Development of peptide inhibitors as molecular therapeutics against tandem-repeat proteins

Rohan Sakariah Eapen

Darwin College
University of Cambridge

This thesis is submitted for the degree of Doctor of Philosophy

December 2019

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. It does not exceed the prescribed word limit for the relevant Degree Committee

Rohan Sakariah Eapen

23 December 2019

Abstract

Tandem-repeat (TR) domains occur in approximately one third of all proteins, yet these domains are less well understood compared with their globular counterparts. TR proteins act as molecular scaffolds that facilitate protein-protein interactions (PPIs) in diverse biological contexts. In disease pathologies such as cancers, misfolded, mutated or deregulated TR function is often the cause. However, targeting tandem-repeat proteins is challenging due to their flat and large interaction interfaces. In this thesis, I have explored the development of peptide-based inhibitors of PPIs involving two targets, Tankyrase (TNKS) the Anaphase Promoting Complex/Cyclosome (APC/C^{Cdc20}).

TNKS belong to the poly-ADP-ribose polymerases (PARP) family of enzymes that PARylate their substrates, leading to ubiquitination and subsequent degradation. TNKS bind their protein substrates through an ankyrin-repeat domain comprising five so-called Ankyrin Repeat Clusters (ARCs). TNKS is of particular therapeutic interest in cancers due to its PARylation of Axin1, a subunit of the β-catenin destruction complex. We investigated the design of peptide inhibitors of TNKS ARC domains through a structure-guided approach. We first attempted to expand the range of biophysical assays to assess ligand binding to TNKS ARCs. We then designed and tested a set of linear and chemically constrained peptides for their ability to bind to the ARCs.

The APC/C is an E3 ubiquitin ligase that has broad substrate-specificity via its two WD40-domain co-activators, Cdc20 and Cdh1, which recognise three peptide motifs (degrons) on its substrates: the D-box, the KEN box and the ABBA motif. The regulation of mitotic exit is tightly controlled by modulating the levels of Cdc20 and Cdh1 through different stages of the cell cycle. Cdc20 overexpression is associated with many cancers. Currently, no specific inhibitors of the APC/C^{Cdc20} exist in the clinic or in clinical trials. We designed a series of synthetic peptides based on the D-box to bind to Cdc20, and thereby modulate APC/C activity. Peptides were tested using surface plasmon resonance and differential scanning fluorimetry assays using recombinant Cdc20. Furthermore, D-box peptides were able to bind exogenous Cdc20 in mammalian cell lysates using a Cellular Thermal Shift Assay (CETSA). Lastly, I determined the co-crystal structures of three of the highest affinity D-box peptides bound to Cdc20, providing high-resolution insights into the binding interface, which should aid in the further development of Cdc20 inhibitors.

"The absence of evidence is not evidence of absence"

Carl Sagan

Acknowledgements

First, I would like to thank my two supervisors, Prof. Laura Itzhaki and Dr. Christopher Phillips who have mentored me through the last four years of my scientific career. I am grateful for their trust and freedom, which allowed me to explore any (and all) projects I found interesting. However, my gratitude extends far beyond their roles as a scientific mentors, as they have provided emotional support whenever the metaphor of living life on the edge came too close for comfort.

I would like to thank all past and present members of the Itzhaki lab for their friendship, shared enthusiasm for science and most importantly late nights in the lab discussing science and life over a whisky or two. Next, much of this work would not have been possible without the opportunities I have received from my mentors and friends in Singapore. In truth, without Dr. Chandra Verma, I would not be where I am today. To Prof. David Lane for allowing me to infiltrate his lab, not once but twice! To Dr. Christopher Brown for instilling and inspiring a sense of scientific rigour and introducing me to the wonderful field of peptide drug discovery. And most recently, to Dr. Fernando Ferrer for taking this 'wannabe' peptide chemist under his wing.

My gratitude extends to the wonderful scientists I have met at AstraZeneca, who have been so open and willing to explore all projects I have thrown their way. In particular, to the 'all-knowing' Dr. Christopher Stubbs who has taken the time to teach me all things biophysics. To Dr. Marianne Schimpl who has inaugurated me into the dark art of crystallography and foretold that the only determinant for getting crystals can be attributed to sheer desperation at the end of your PhD!

To all the wonderful people of the rowing world, whom are too numerous to mention here. You have kept my spirits high throughout the past four years, come rain or shine on the winding river Cam.

To my dearest family, who have always been by my side through the ups and downs of life over the past 26 years. And lastly, to Marie. You have made the last three years of life the most exciting yet. As the Cambridge chapter in our lives comes to a close, I can't wait to see what the next chapter will bring.

Contents

Copyright

List of Figures

List of Tables

List of Abbreviations

^{1}H NMR	Proton Nuclear Magnetic Resonance
2xYT	2x Yeast-Tryptone
3BP2	Abl SH3-binding protein 2
ADP	Adenine Diphosphate
AKTA	Trade Name for purification system from GE Healthcare
YFP	Am Yellow Fluorescent protein
ANK	Ankyrin
APC	Adenomatous polyposis coli
Apc	Anaphase promoting complex subunit
APC/C	Anaphase Promoting Complex / Cyclosome
APS	Ammonium Persulphate
ARC	Ankyrin Repeat Clusters
ATP	Adenine Triphosphate
BCA	Bicinchoninic Acid
BEH	Ethylene Bridged Hybrid
BEVS	Baculovirus Expression Vector Systems
BRD4	Bromodomain-containing protein 4
BSA	Bovine Serum Albumin
C_3	(*S*)-2-amino-4,4-dimethylpentanoic acid
CAG	Cytomegalovirus, chicken beta-actin and rabbit beta-globin
CAR-T	Chimeric Antigen Receptor T-cells
cBuAla	(*S*)-2-amino-3-cyclobutylpropanoic acid
CCP4i	Collaborative Computational Project No. 4
Cdc20	Cell division cycle 20
Cdh1	Cell division cycle 20 homolog 1
Cdk1	Cyclin-dependant kinase 1
CETSA	Cellular Thermal Shift Assay
Cl_3	(*S*)-2-amino-4,4,4- trichlorobutanoic acid
CMV	Human Cytomegalovirus
CNS	Central nervous system

Coot	Crystallographic Object-Oriented Toolkit
CPPs	Cell Penetrating Peptides
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
Cryo-EM	Cryogenic Electron Microscopy
CST	Cell Signalling Technologies
CTPR	Consensus Tetratricopeptide Repeat
CuAAC	Cu(I)-Catalysed Azide-Alkyne Cycloadditions
CV	Column Volumes
DARPins	Designed Ankyrin Repeat Proteins
D-box	Destruction-Box
DCE	Dichloroethane
ddH_2O	Double Distilled Water
DELs	DNA-Encoded Libraries
DIEA	*N*,*N*-Diisopropylethylamine
DMEM	Dulbecco's modified eagle medium
DMF	Dimethylformamide
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
DNase I	Deoxyribonuclease 1
DSF	Differential Scanning Fluorimetry
DTT	Dithiothreitol
E.coli	*Escherichia coli*
ECL	Enhanced chemiluminescence
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDTA	Ethylenediaminetetraacetic acid
eq.	Equivalent
ESI-MS	Electrospray Ionisation Mass Spectrometry
Et_2O	Diethylether
ExPASy	Expert Protein Analysis System
F_3	(*S*)-2-amino-4,4,4-trifluorobutanoic acid
FBS	Fetal Bovine Serum
Fc	Flow Channel
FCS	Fetal Calf Serum
FD	FastDigest
FDA	Food and Drug Administration
Fmoc	9-fluorenylmethoxycarbonyl
FP	Fluorescence Polarisation
GPCRs	G-Protein Coupled Receptors

GS	glycine-serine
GST	Glutathione S-Transferase
HATU	Hexafluorophosphonate azabenzotriazole tetramethyl uronium
HCl	Hydrochloric Acid
HEK 293T	Human embryonic kidney 293 expressing SV40 large T antigen
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
His_6	Hexahistidine
HIV-1	Human Immunodeficiency Virus
HOAt	1-Hydroxy-7-azabenzotriazole
HPLC	High performance Liquid Chromatography
HPS	Histidine Proline Serine
HRP	Horse Radish Peroxidase
HSA	Human Serum Albumin
HSB	High Salt Buffer
HTS	High Throughput Screening
IC_{50}	Half maximal Inhibitory Concentration
IDR	Intrinsically Disordered Region
IEX	Ion Exchange
IMAC	Immobilised Metal Ion Affinity Chromatogrphy
IPA	Isopropanol
IPTG	Isopropyl-β-D-thiogalactoside
IRAP	Insulin-regulated aminopeptidase
IR-tail	Isoleucine-Arginine tail motif
ITC	Isothermal Titration Calorimetry
KEN-box	Lysine-Glutamate-Asparagine-box
LB	Lysogeny Broth
LCMS	Liquid Chromatography Mass Spectrometry
LLC	Limited Liability Company
loxP	locus of X-over P1
LSB	Low Salt Buffer
m/z	Mass to charge ratio
Mad2	Mitotic arrest deficient 2
Mad3	Mitotic arrest deficient 3
MBP	Maltose-binding protein
MCC	Mitotic Checkpoint Complex
MD	Molecular Dynamics
Mdm2	Mouse double minute 2
MeCN	Acetonitrile

MeOH	Methanol
MES	4-Morpholineethanesulfonic acid
MOPS	3-(N-morpholino)propanesulfonic acid
MPD	2-Methyl-2,4-pentanediol
MQ	MilliQ
mRNA	Messenger Ribonucleic Acid
mTOR	mammalian Target of Rapamycin
M_w	Molecular Weight
MWCO	Molecular Weight Cut Off
NAD^+	Nicotinamide Adenine Dinucleotide
NHS	N-Hydroxysuccinimide
Ni-NTA	Nickel-Nitrilotriacetic acid
NMP	N-Methyl-2-pyrrolidone
NP-40	Nonidet-P40
NuMA	Nuclear Mitotic Apparatus Protein 1
OD_{600}	Optical Density at 600nm
ORF	Open Reading Frame
OS	Operating System
PAR	Poly-ADP-Ribose
PARP	Poly-ADP-Ribose Polymerase
PARP1	Poly-ADP-Ribose Polymerase 1
PBS	Phosphate Buffered Saline
PBST	Phosphate Buffered Saline Tween20
PC	Proteasome/Cyclosome repeat
PCR	Polymerase Chain Reaction
PDA	Photo Detector Array
PDB	Protein Data Bank
PEG	Polyethylene Glycol
PES	Polyethersulfone
pI	Isoelectric point
pK_a	Acid dissociation constant
PMSF	Phenylmethanesulfonyl Fluoride
PNAC	Protein and Nucleic Acid Chemistry
PNK	T4 Polynucleotide Kinase
PPIs	Protein-Protein Interactions
PROTAC	Proteolysis Targeting Chimera
psi	Pounds per Square Inch
PTFE	Polytetrafluoroethylene

PVDF	Polyvinylidene Difluoride
QTOF	Quadrupole Time of Flight
R_5	(*R*)-2-(4'-pentenyl) alanine
R_8	(*R*)-2-(7'-octenyl) alanine
RCM	Ring-closing Olefin Metathesis
RCSB	Research Collaboratory for Structural Bioinformatics
RE	Restriction Endonuclease
RI	Refractive Index
RING	Really Interesting New Gene
rmsd	root mean squared deviation
RNA	Ribonucleic acid
RNase A	Ribonuclease A
RNF146	RING finger protein 146
rpm	revolutions per minute
RTH	Round the Horn
RU	Response Units
S_5	(*S*)-2-(4'-pentenyl) alanine
SA	Streptavidin
SAC	Spindle Assembly Checkpoint
SAM	Sterile Alpha Motif
SAR	Structure-activity relationship
SAXS	Small-angle X-ray Scattering
SBDD	Structure-Based Drug Design
SDM	Site Directed Mutagenesis
SDS	Sodium Dodecyl Sulfate
SDS-PAGE	Sodium dodecyl Sulphate Polyacrylamide Gel Electrophoresis
SE	Standard Error
SEC	Size Exclusion Chromatography
SEM	Standard Error of the Mean
SLiM	Short Linear Motif
SOC	Super optimal broth
SPPS	Solid Phase Peptide Synthesis
SPR	Surface Plasmon Resonance
TAE	Tris-Acetate-EDTA
TAD	Transactivation domain
TAME	Nα-p-Tosyl-L-arginine methyl ester
TAMRA	Tetramethylrhodamine
TAT	Trans-acting Activator of Transcription

t-BuOH	tertiary butanol
TCEP	Tris(2-carboxyethyl)phosphine hydrochloride
TEMED	Tetramethylethylenediamine
TEV	Tobacco Etch Virus
TFA	Trifluoroacetic acid
TfbI	Transformation Buffer I
TfbII	Transformation buffer II
TGS	Tris-Glycine-SDS
THPTA	Tris(3-hydroxypropyltriazolylmethyl)amine
TIPS	Triisopropylsilane
TLM	Translocation Motif
T_m	Melting Temperature
TNKS	Tankyrase
TPR	Tetratricopeptide Repeat
UPLC	Ultra Performance Liquid Chromatography
UV	Ultra Violet
UVP	Ultra-Violet Products
VHL	Von-Hippel Lindau
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Chapter 1

General introduction

1.1 Drug discovery approaches

Drug discovery is a broad and diverse field tackling a wide range of disease pathologies, such as oncology, cardiovascular diseases and microbial infections, only to name a few. Each disease pathology brings its own unique set of challenges, to which the drug discovery field must constantly adapt and accommodate accordingly. The emergence of new technologies, such as chimeric antigen receptor T-cells (CAR-T) are incorporated into new therapeutics, which open largely unexplored avenues of harnessing the patient's immune system to treat the disease [1]. Another promising technology that has recently caused ethical concerns, is the use of clustered regularly interspaced short palindromic repeats (CRISPR), in which the possibility of eradicating the disease pathology can be tackled at the gene level [2]. Despite the ongoing development of these promising techniques, the vast majority of successful drugs that have made it into the clinic are small molecules or larger biologics such as antibodies. In this section, the aim is provide a perspective of the difficulties and challenges facing drug discovery and the more common modalities that are typically explored.

1.1.1 A historical perspective of drug discovery

The earliest forms of drug discovery or rather the use of compounds as medicines to treat ailments date back over 4000 years. Accounts of apothecaries with lists of ailments and their remedies, define the first use of a modern-day prescription with the ancient Mesopotamians, Babylonians, Egyptians, Greeks, Indians and Chinese [3]. In these times, remedies often took the form of natural products, such as herbs, roots and spices in ground forms mixed in a suitable liquid medium, often wine or beer. Of course, the medical and pharmaceutical field has significantly advanced since these times, but more major advances were only prevalent in the last 150-200 years.

Unfortunately for human society, major technological and medical advancements have been made as a result of international war. The most notable example of advancement in drug discovery arguably revolves around antibiotics in the early 20th century and their pivotal use in saving the lives of soldier during World War II for treating infectious disease that were often acquired during life-saving amputations or surgery [4]. Alexander Fleming's most discovery of penicillin in 1928 was a landmark in of the development of specific compounds that could be used to treat disease [5]. However despite it's discovery, it took several more years till it was widely used in humans [4]. The era of antibiotics and in particular the search for new and more potent antibiotics to target gram-positive and gram-negative bacteria sparked the evolution of more designed screening approaches to be used. The study led by the then graduate student, Albert Schatz and his supervisor Professor Selman Waksman was a landmark in systematic screening of antagonistic *Actinomycetes* organisms that produced a growth inhibitory effect against a variety of bacteria, but most notably against *Mycobacterium tuberculosis* [6]. The discovery of this potent molecule, streptomycin, later resulted in the awarding of the Nobel Prize in Physiology or Medicine in 1952 to Prof. Waksman. However, perhaps more influential than the discovery of streptomycin itself was the concept of systematically screening molecules (or rather organisms, in this case) against the cause of the pathology or indeed varying extents of models to describe or replicate the underlying pathology or its cause. This concept paved the way for high-throughput screening, in what is now considered standard in modern day pharmaceutical drug discovery.

Evolution of drug regulation

It should be noted that in the last century, the growth of scientific knowledge has increased at an incredible rate. As such, while medicines discovered early in this period do work, the mechanisms by which they elicited their functional effect were not well understood. Regulatory bodies, such as the United States Food and Drug Administration (FDA) are in place to decide whether a new drug molecule is safe for use and elicits the desired/ expected phenotype. Whilst many drugs were deemed safe, and thus granted approval for use in the clinic, this regulation was not sufficient in the case of thalidomide and marked a turning point in scientific and clinical rigour for the safety of patients. The drug molecule was found to alleviate morning sickness symptoms in pregnant mothers, however the advent of severe defects in offspring provided a important lesson to the drug industry and its regulators of how impactful wrong decisions can be [7]. Since the advent of Thalidomide, the FDA and other regulatory bodies now require significantly more information with the introduction of the Kefauver-Harris Drug Amendments, and studies to define efficacy and safety of the proposed molecule are now required [8].

Whilst this was undoubtedly a move into the correct direction, the impact was significant given how limited knowledge in the scientific community of biology at the molecular level was (and perhaps still is). To provide some perspective of how limited our knowledge was, the idea that DNA is transcribed to RNA, which then is translated into proteins had just been suggested by Francis Crick in September of 1957, at University College London [9]. Only one year preceeding this revolutionary concept, was Thalidomide in the clinic. Coupled with the lack of knowledge at the time of what are now considered the core fundamentals of cell and molecular biology, the true complexities of drug action can also be unclear. Despite the use of Thalidomide from 1956 to 1961 worldwide, it's biological effector remained elusive for many years. Even during this time, attempts to repurpose the molecule in the treatment of Leprosy (which was later discouraged by the World Health Organisation) and cancers with solid tumours lacked information on the molecular interactions of the molecule. In cancers, it was used for its anti-angiogenic properties as an inhibitor of basic fibroblast growth factor-induced angiogenesis. Only in 2010, over 50 years since it's initial introduction as a drug molecule for pregnant women, did we learn that the molecular target of Thalidomide was Cereblon [10]. This raises an important question for the drug discovery community. Before a drug is used in humans, how much should we know about the drug's molecular interactions, downstream effects? What are suitable models to determine safety and efficacy? If a molecule produces the desired phenotype or therapeutic action, but the mechanism of action is unclear, should it be used?

1.1.2 The growing problem in drug discovery

Fast forward from the early days of drug discovery to present day, the scientific field has grown enormously. Despite this, the range of targets and success of projects pursued by pharmaceutical companies is very limited. This owes to multiple and complex issues, which often intertwine very heavily with each other. In this section, I will try briefly outline some of the challenges the industry faces.

Is pharma only interested in profits?

One would like to believe that everyone involved in drug development from the research scientists performing experiments up to group leaders and heads of departments directing research and therapeutic areas and all the way to business leaders and chief executive officers that everyone's primary goal is to make a positive impact to patients. While this might true, it would be naïve to assume that a business can survive without being minimally financially sustainable. The largest and most established Pharmaceutical com-

panies, such as AstraZeneca, reported a total revenue values of USD $24.4 bn [11] for 2019. Despite this, the money going back directly into R&D is only USD $6.059 bn [12], approximately 25%. This raises the question, would better outcomes, i.e. more drugs be made if more money was dedicated to R&D? Interestingly, significantly more money (inflation-adjusted) has been made available towards research and development purposes per approved drug over the years from 1950 to 2010, yet the number of drugs approved as time has progressed has declined [13, 14]. It would therefore be illogical to think that a lack of R&D funding is entirely responsible for this and the inefficiencies of bringing new drugs to market must come from elsewhere.

How good is the model?

Models form a crucial element to the determination of whether a compound produces the desired outcome. The definition of an outcome can be many things. Does it bind to the target? Does it produce an inhibitory effect? Is there a phenotypic change? Does the patient see an improvement in their disease? As knowledge has expanded over the past century, the scientific community has been able to further dissect down and simplify 'the model' from the *in vivo* setting observing the final outcome towards more focused and individual target oriented models *in vitro*. It can be argued that there are positives and negatives from both perspectives and one must also take into account the feasibility of doing so by taking into account the current technologies, knowledge and ethical considerations.

If one considers the models originally used by Schatz et al. [6], where they screened *Actinomycetes* strains that produced an inhibition of bacterial growth, this would class as looking at the final output, i.e. did the strain kill the pathogenic bacteria? While this is considered ethically sound, doing similar types of experiments in mice or even humans with the large libraries of small molecule compounds, natural products or biologics that are available today would not be acceptable on an ethical standpoint. Although, it would be likely to find suitable candidates that produce the desired therapeutic effect, the sheer numbers of negatives that would result in severe toxicity or even death make this approach immoral. This has forced the scientific community to 'simplify' the model from the *in vivo* setting towards cellular models and even further towards examining the nature of the ligand-receptor interaction.

The adoption of *in vitro* models has advantages, but also a major disadvantage. Firstly, these simplified models provide a means to screen a larger number of compounds in a more quantitative manner to provide an output of potential downstream success. These can be done in highly controlled systems utilising many different models from assessing binding

of the binary components, such as the ligand and the receptor in a purified system, but also in more complex cellular models. However, with all of these models, they do not take into account issues of dealing with a whole organism. This introduces several unknowns, such as immune response, tissue distribution, metabolism of the drug molecule and clearance from the organism to name but a few. Even then, at the *in vivo* level, is the model organism itself correct? This can be a cause of concern as mice, dogs or monkeys may not fully recapitulate the homology, complexities and intricacies for a given disease phenotype [15, 16]. This leaves the drug discovery community in a state where the current standard of *in vitro* and *in vivo* models are the only option for perform drug discovery, without further innovation in developing more robust models that recapitulate the human physiology and biology in its fullest extent.

Targeting what is easy

Pharmaceutical companies, in particular have a tendency to target the so-called 'easy targets'. Easy, in this context doesn't necessarily mean that the disease phenotype is easy to overcome, but rather is an indication of the ease of developing a molecule to satisfy one or several of the models developed for the disease pathology of interest. Kinases, GPCRs and cell surface receptors form the majority of protein classes targeted in pharmaceutical drug discovery research programs. Kinases, in particular are targeted since the role of phosphorylation in the regulation of proteins is very central in the regulation of biological activity [17]. The addition of a phosphate is produced by the hydrolysis of ATP or GTP (in the case of GPCRs) to ADP and GDP, respectively, and thus these molecules invariably form binding interactions in the active site of the enzyme. Therefore, there are varying degrees of structural homology around the active sites of all kinases, which in turn can be exploited for the development of inhibitors against these pockets. It is also no surprise then, that pharmaceutical companies utilise kinases as targets for potential cancer therapeutics when phosphorylation of a various proteins is fundamental in driving cell cycle progression, (reviewed by Bhullar et al. [18] and references therein). With the evolution of targeting proteins sharing conserved or structurally homologous active sites, it is no wonder that the vast majority of molecules found in small molecule libraries that are designed and synthesised by medicinal chemists are bias towards targeting these particular protein families [13, 14, 19].

A fundamental process of drug discovery, is the process of understanding the structural basis of the protein-ligand interaction. For this to be understood, the protein must be structured. However, while many proteins contain ordered regions, approximately 44% also contain regions (classified as greater than 30 amino acids) that form no defined secondary structural elements [20, 21] and these intrinsically disordered regions (IDRs) play a

large and diverse number of roles regulating biological function. The typical response from a drug discovery perspective is to ignore these 'difficult' targets. However, these difficult targets can often be of valuable therapeutic interest. To provide direct examples of these, in relation to work conducted in this thesis and collaborative work, I will highlight two such examples. The first is p53, whose transactivation domain (TAD) comprising residues 1-93 is reported to be intrinsically disordered [22, 23]. The TAD domain is able to form numerous interactions with proteins such as Mdm2, MdmX, Taz2 and CBP/p300, among others [24]. It's ability to interact with so many proteins places it towards the centre of cellular signalling pathways and its' intrinsic disorder is crucial for this. Stabilising p53, the so-called the 'guardian of the genome' [**?**] through blocking Mdm2/MdmX-mediated ubiqutination and subsequent degradation has long been known as a important therapeutic avenue. Efforts to stabilise wild-type or mutant p53, particularly in its intrinsically disordered regions are far less explored, but are nonetheless valid approaches [25, 26]. The second example is that of Cdc20 and Cdh1, of which the former is the greater focus of this thesis. Briefly, Cdc20 and Cdh1 interact with the core Anaphase Promoting Complex/ Cyclosome (APC/C) through intrinsic disordered regions, which flank the core WD40 domain. Without these regions, functional output of both proteins is severely attenuated [27, 28]. Particularly in the case of Cdc20, attenuating its function in oncology is of valuable interest and therefore is a potential avenue to be explored. It should be noted that a more in-depth analysis of this protein is discussed later in this chapter and further in Chapter 4, Figure 4.1.

The sections above are by no means intended to be a fully comprehensive overview of the current issues in drug discovery. However, they do provide some insights that don't often reach the public and perhaps even academic readership. Having identified some of the issues involved in the drug discovery pipeline, the more pertinent questions and issue to address is how to change the industry to improve the number of successful drugs in the clinic to positively impact the lives of patients? This is of course a multi-faceted challenge and will require input not only from the pharmaceutical industry, but also academia and biotech who play a large role in the development of new technologies and introduction of new knowledge and principles that can be applied within the pharmaceutical industry.

1.1.3 Small molecule drug discovery

Small molecule are the most favoured drug class for intracellular targets and G-protein coupled receptors (GPCRs). They typically target protein-ligand interactions, such as the active sites of kinases and therefore are often developed against larger pockets within the target. In the pharmaceutical industry, this makes targeting protein-ligand interactions

more favourable than targeting protein-protein interactions (PPIs), which often do not have well defined pockets within their binding interfaces [29].

High-throughput screening (HTS) campaigns using large libraries (>1 $\times$ 10^9 compounds) of small molecules for screening in functional assays to identify molecules that produce a desired phenotype. Hits identified in these screens need to be further validated by biophysical methods (reviewed in [30]). These libraries can yield a number of hits for targets such as kinases, yet the success rate of inhibiting PPIs is much lower [29, 31]. A major reason for the low success rate is the chemical and spatial diversity libraries favouring the typically explored protein classes [19, 32, 33]. Additionally, PPI interfaces are highly diverse in their general topology (Figure 1.1), which adds to the complexity of generating small molecule libraries that can efficiently explore all of this space. Lastly, these interfaces typically cover a larger surface area (1500-3000 $Å^2$) versus the surface area of small molecule binding sites (300-1500 $Å^2$) [30]. Despite these challenges, multiple small molecule inhibitors against PPIs exist, such as Nutlin-3a for the Mdm2-p53 interaction [34] and (+)-JQ1 for the BRD4-H3K13ac interaction [35]. However, developing potent ligands to target PPIs often necessitates a fragment-based approach, in which multiple low-affinity fragments are identified and then 'stitched' together to generate a high-affinity ligand [36, 37]. While these methods do provide ligands, their efficacy downstream might be limited due to the smaller interaction surface compared to the native PPI and thus transient or limited activity due to the native PPI may still occur [38]. In order to improve functional efficacy, larger molecules, such as macrocycles and peptides offer suitable alternatives.

1.1.4 Macrocycles

The term 'macrocycles' encompasses a broad range of compounds of both natural and synthetic origin [39, 40]. This class of molecule bridges the gap between small molecule and larger biological therapeutic molecules, such as peptides mini-proteins and antibodies. Their increased structural complexity allows the targeting of larger interfaces found in PPIs [41]. However, this poses a challenge to rational design of macrocyclic compounds, especially in the absence of a starting structure, and hence complicates structure-based drug design (SBDD). As a result, many synthetic macrocyclic compounds are produced by library screening methods and are further optimised by SBDD [42, 43]. Without the use of libraries, the discovery of macrocycle PPI inhibitors is a difficult task, unless the design in based on existing natural macrocycles such as Rapamycin, which targets the mammalian Target of Rapamycin (mTOR) [44].

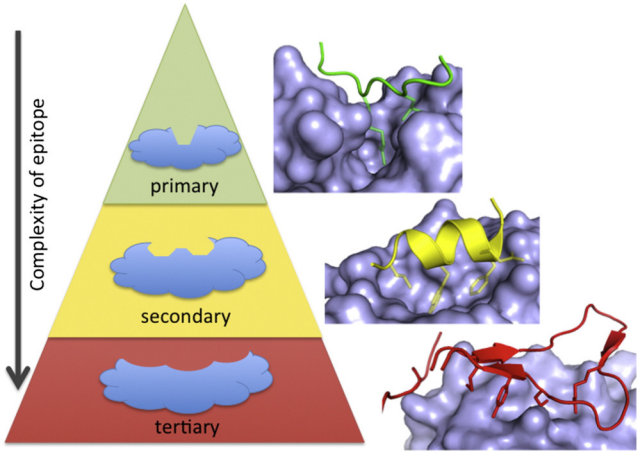


Figure 1.1: Protein-protein interactions vary in their chemical and structural features. PPIs can be formed by primary amino acid sequences without defined secondary structures, also known as Short Linear Motifs (SLiMs). Formation of secondary structure elements such as α-helices or β-sheet structures further increases the complexity of the interaction. Tertiary interactions are formed by elements of primary and/or secondary structures in the context of a protein, e.g. antibodies. Figure adapted from Arkin *et al.*. [31]

.

1.1.5 Peptides

The use of peptides as therapeutic agents originated with insulin being used to treat patients suffering from diabetes mellitus [45]. Insulin itself, is composed of two polypeptide chains, A (21 amino acids) and B (30 amino acids) that are cross-linked by two intermolecular disulfide bonds. Early extractions of insulin for use in humans were sourced from fetal calf pancreas, though later this would be extracted from porcine source which is more structurally homologous to human insulin with only one amino acid difference. Several other natural peptide hormones were gradually introduced over the years, such as the Adrenocorticotropic hormone [46], Calcitonin [47] and Oxytocin [48]. The field of peptide therapeutics has largely revolved around targeting extracellular binding sites, such as those on GPCRs (Figure 1.2) [38]. Most notably, and relevant to aspects of the work presented in this thesis, intracellular targets make a very small proportion of those that have entered clinical trials. One example of such a peptide is the p28 peptide, which has been indicated for used in p53(+) patients with solid tumours and central nervous system (CNS) tumours [49, 50]. These studies alongside others, such as Aileron's Mdm2 and MdmX dual stapled peptide inhibitor, ATSP-6924 [51–54], highlight that intracellular targets for peptide modalities is a viable option, that may have success in the clinic.

Exploring intracellular targets using natural peptides, unlike macrocycles or small

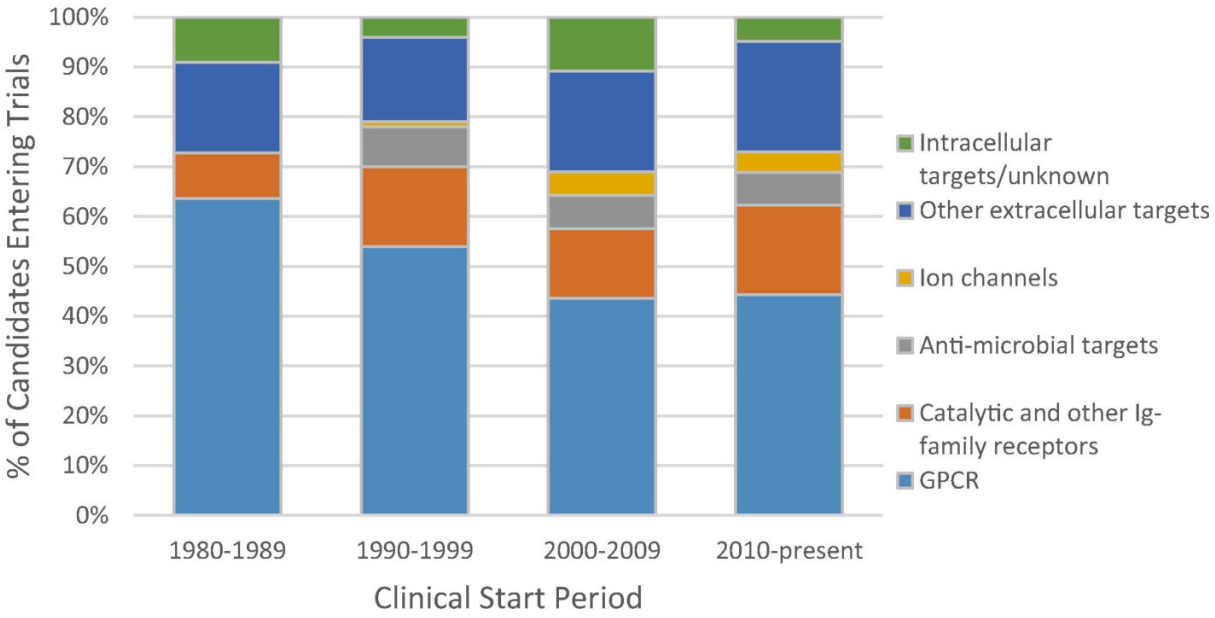


Figure 1.2: Distribution of targets for peptides entering clinical trials over the years. Figure adapted from J.L. Lau and M.K. Dunn [38].

molecules, is more challenging to develop efficient screening methods due to their lack of cell permeability (further discussed later). To identify binding motifs, amino acid sequences that are responsible for mediating PPIs are often determined by mutagenesis studies, e.g deletion and/or alanine substitution of specific residues [55, 56]. This approach is limited in its effectiveness and may not fully capture the complexity of the binding interaction. In addition, biology has not necessarily evolved to produce high-affinity PPIs, but has instead designed transient and weak affinity PPIs to mediate complex biological pathways [57, 58]. Therefore, it is quite common that peptides derived directly from a protein-protein interaction do not consist of the optimal amino acid sequence for high affinity binding.

Similar to small molecules and macrocycles, the generation of high-affinity peptides against a target is often performed by library screening in a cell free environment. Several library screening approaches have been developed, such as Phage [59, 60], mRNA [61] and ribosome display [62]. While these techniques often produce a high-affinity peptides, diverse libraries of the required peptide length are not always available or indeed accessible. In addition, not all PPIs can be targeted using these methods. For example, if the protein of interest is part of a larger macromolecular assembly, valid targeting approaches may have to consider the complex as a whole.

1.2 Challenges of peptides as therapeutics

1.2.1 Cell permeability

Lipinksi's 'rule of 5' defines the ideal characteristics of a drug molecule and includes parameters for the molecular weight ($\leq$ 500 Da), logP (logarithm of 1-octanol/water partition coefficient, $\leq$ 5), number of H-bond donors ($\leq$ 5) and H-bond acceptors ($\leq$ 10) [63]. Although these rules are gradually being broken down as new and existing technologies improve, they still form a major basis of drug discovery at present and therefore pose a challenge to the successful development of peptide inhibitors, which rarely adhere to these rules. Diffuse cell permeability of designed peptides is rarely observed unless specifically selected for. During the past decade, Lokey and co-workers have investigated passive cell permeability of cyclic pentapeptides and hexapeptides [64–67]. All natural and designed cyclic peptides were able to traverse the cell membrane if their estimated AlogP values were between between 3-5, thus adhering to Lipinski's rule of 5 [68, 69]. At present, these rules have not been applied to linear peptides motifs, possibly due to their increasing conformational complexity.

In efforts to deliver peptide and larger biologic therapeutic molecules intracellularly, several different techniques have been explored. Nanoparticle-mediated and liposomal delivery methods have been explored extensively for the delivery of small molecules, peptides and biologics (reviewed in [70, 71] and references therein). However, these methods require extensive expertise in formulation of the specific delivery vehicle for each respective drug, and therefore are not ideal for the identification of positive hits early in the screening process. However, once an active peptide or biologic is found, nanoparticles offer an attractive route of administration. Typically biologics and peptides would have to be administered through the intravenous route. This is generally not favourable since administration will always have to be performed by a medical professional. Thus, orally availability of drug molecules is often preferred. Nanoparticle offer this administration route due to their ability to traverse the gastrointestinal tract via endocytosis mechanisms [72]. The physical barrier of the nanoparticle provides shielding of the cargo from excreted proteases. These properties can be extensively modified through the variation of nanoparticle material, diameter, surface charges and modification.

1.2.2 Cell penetrating peptides

An alternative approach to intracellular delivery is the use of, cell penetrating peptides (CPPs) to mediate the entry of a cargo molecule (reviewed in [73, 74]). Early examples of polycationic CPPs in the literature were derived from trans-acting activator of transcription (TAT) protein from human immunodeficiency virus (HIV-1). A short sequence derived from this protein (YGRKKRRQRRR) was shown to be sufficient for cell entry in mouse fibroblasts [75, 76]. Another commonly adopted sequence is the amphipathic penetratin CPP (RQIKIWFQNRRMKWKK) [77, 78]. In addition, poly(R) peptides such as R_9 (RRRRRRRRR) [79], and many others (see Table 1.1) have been used. Interestingly, amphipathic CPPs appear to contain cationic residues at positions i,i+3, i,i+4, i,i+7 and i,i+11, which correspond to one face of an α-helix. Two studies have linked the relative position of cationic residues and the helix-forming propensity of CPPs to cellular uptake efficiency [80, 81].

Table 1.1: List of common cell penetrating peptides. Cationic residues are coloured in red. Table adapted from Fonseca *et al.*. [73].

Cell-penetrating peptide	Amino acid sequence
TAT_{49-57}	RKKRRQRRR
Polyarginine	RRRRRRRRR
R_9F_2	RRRRRRRRRFF
Decalysine	KKKKKKKKKK
Penetratin	RQIKIWFQNRRMKWKK
Transportan	GWTLNSAGYLLGKINLKALAALAKKIL
HIV-Tat derived PTD4	YARAAARQARA
Hepatitis B Virus Translocation Motif (TLM)	PLSSIFSRIGDP
$mPrP_{1-28}$	MANLGYWLLALFVTMWTDVGLCKKRPKP
POD	$GGG(ARKKAAKA)_4$
pVEC	LLIILRRRRIRKQAHAHSK
ARF_{1-22}	MVRRFLVTLRIRRACGPPRVRV
EB1	LIRLWSHLIHIWFQNRRLKWKKK
Rath	TPWWRLWTKWHHKRRDLPRKPE
CADY	GLWRALWRLLRSLWRLLWRA
Histatin 5	DSHAKRHHGYKRKFHEKHHSHRGY

1.2.3 *In vivo* half-life

Peptides are intrinsically subject to proteolytic cleavage at amide bonds, which often leads to short *in vivo* half lives. This severely limits the administration route, whereby peptides delivered orally to a patient are subject to the acidic conditions and peptidase-rich environment of the digestive system. Additionally, uptake of peptides through the gastrointestinal tract is particularly poor due to poor penetration of the mucosal membrane [82]. To begin addressing these issues in early stage drug discovery, researchers have adopted multiple strategies applicable to both peptide and protein therapeutics. The addition of D-amino acids at the N- and C- terminals of a peptide can improve the half-life of the peptide in serum [83, 84]. These can also be incorporated at cleavage 'hotspots' to further protect peptides from proteolysis. Furthermore, 'retro-inverso' peptides, in which the entire L-peptide is substituted for a mirrored D-amino acid-containing sequence, have been shown to resist proteolysis [85, 86]. However, as noted by Li *et al.*. these peptides do not always yield inhibitors that are able to bind the target protein [87].

Another commonly explored technique within the biologics field is the use of human serum albumin (HSA)-targeting moieties, which serve to increase the plasma half-life of the biological agent (reviewed extensively in [88]). This strategy has been effective for different classes of biologics such as DARPins [89] and bicyclic peptides [90], which would otherwise be cleared rapidly *in vivo*. A further example, includes the use of fatty acid chains to improve the serum half life of the glucagon-like peptide-1 receptor agonist, Liraglutide [38, 91]. The C16 fatty acid, incorporated from a lysine sidechain is able to bind to HSA at one of its seven fatty acid binding sites (Figure 1.3) [92], therefore increasing its retention in the plasma and decreasing its clearance.

1.2.4 Macrocyclic peptides

Macrocyclisation, also commonly known as stapling or constraining, in the context of a peptide refers to the bridging of two or more amino acids with a chemical moiety to conform and constrain the peptide towards its bioactive conformation. Compared to other peptide modifications, macrocyclisation has now formed its own field of research and includes a variety of different cyclisation strategies (reviewed in [93]). Macrocyclisation is generally categorised into one-component (two reactive groups within one molecule) or two-component techniques (two or more molecules that react with each other). Sidechain lactamisation [94], Cu(I)-catalysed azide-alkyne cycloadditions (CuAAC) [95, 96], and ring-closing olefin metathesis (RCM) [97] can be used in one-component macrocyclisation techniques. However, these reactions also form the basis of two (or more)-component

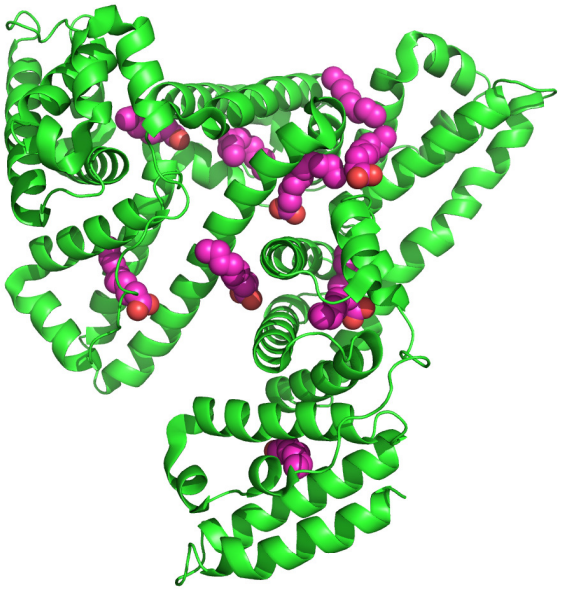


Figure 1.3: Fatty acid binding sites of human serum albumin (HSA). HSA shown cartoon as cartoon representation in green and decanoic acid shown as spheres in magenta. PDB: 1E7E [92].

techniques, thereby increasing the chemical space that can be explored [98–101]. Each of these strategies provide different physical and geometrical properties to the constraining macrocycle and it is not necessarily clear which strategy is best for a given PPI system, particularly in the context of non-helical peptides.

One major advantage of these macrocyclisation strategies is the increased *in vitro* stability of the constrained peptide versus the native or non-cyclised peptide [102, 103]. Resistance to proteolysis is likely due to the shielding of cleavable amide bonds by the staple itself and/or the induced conformation. In the context of p53-hydrocarbon stapled peptides targeting Mdm2/MdmX, the macrocycle has a two-fold effect. On the one hand, the peptides are resistant to proteolysis and the other hand the staple mediates cell permeability [104–106]. The uptake mechanism of these peptides is thought to involve endocytotic pathways, although the exact mechanism is not fully understood. It is believed that the aliphatic hydrocarbon staple is crucial for uptake, since non-metathesised peptides are not cell permeable (Dr. Christopher Brown, A*STAR, personal communications). The success of macrocyclisation strategies in peptide therapeutics is highlighted by the development of stapled peptides that targets the interaction of Mdm2/MdmX and p53, ATSP-6924 (Aileron therapeutics), which is currently in Phase 1b/2a clinical trials for the treatment of small cell lung cancer and solid tumours [51–54].

1.2.5 The future of peptide therapies

Targeting protein-protein interactions provides a significant challenge due to the complex nature of the interactions and their regulation. Although small molecules often fulfil the category of cell permeability and target binding, their smaller binding interfaces and therefore antagonism of the problematic interaction can be limited and the molecule is therefore not efficacious. In addition, finding suitable molecules in small molecule libraries, which contain bias towards protein-ligand interactions often results in poor quality hits. While fragment-based drug design offers an alternative to tackling this issue, it is by no means simple to do so and the protein systems must be amenable to the methods used for these studies, such as X-ray crystallography and NMR, which is not always feasible for a given target.

Peptides, with their larger interfaces, compared to small molecules mitigate this initial issue. However, the many challenges that still face small molecules of affinity, efficacy, tissue targeting and half-life to name a few are still relevant to peptides. However, I strongly believe that as new technologies are introduced and are applied within the peptide therapeutic field such as those already mentioned in previous sections, will drastically improve the number of intracellular and extracellular targets that peptide therapies can be applied to. While it may still be a number of years till we see a breakthrough that can be applied more universally, I have no doubt that it will come.

1.3 Tandem-repeat proteins in oncology

Repeat proteins comprise approximately one-third of all proteins within the genome [107, 108]. Solenoids and toroids are two major subclasses, defined by their open-ended and closed structures, respectively, which are composed of multiple repeats of small structural units (typically 20-40 residues repeated multiple times in tandem) [109, 110]. Members of these two classes have a multitude of functions within organisms, including scaffolds of catalytic complexes (e.g. PR65 [111] and Tankyrases [112]) and transcription factors (e.g. β-catenin [113, 114]). However, in many disease states, the deregulation of these repeat proteins causes an increase in the expression of oncogenes [113]. In contrast, overexpression of the ankyrin-repeat protein, Gankyrin, leads to the increased degradation of the retinoblastoma and p53 tumour suppressor proteins [115–117]. For many of these tandem-repeat proteins, small molecule therapeutics are unlikely to yield suitable drug candidates due to the intrinsically flat surfaces through which these proteins interact with their binding partners. The remainder of this chapter will focus on two tandem-repeat proteins involved in oncogenic processes, Tankyrase and the APC/C^{Cdc20}.

1.4 Tankyrase

Tankyrases (TNKS) belong to a family of enzymes known as poly-ADP-ribose polymerases (PARPs). Small molecule inhibitors are readily explored due to the intrinsic 'druggability' of these active sites. However, targeting PARP domains comes at a cost of specificity leading to off-target activity due to the similarity of ligand binding sites between family members [112]. It is also clear that normal function and disease pathology of some PARP proteins can be independent of PARP function and so there is a need to target these proteins through an alternative strategy. Aberrant TNKS activity is implicated in many disease states, ranging from cancer [118–120] and diabetes [121, 122] to rarer disorders such as cherubism [123, 124].

1.4.1 Tankyrase structure

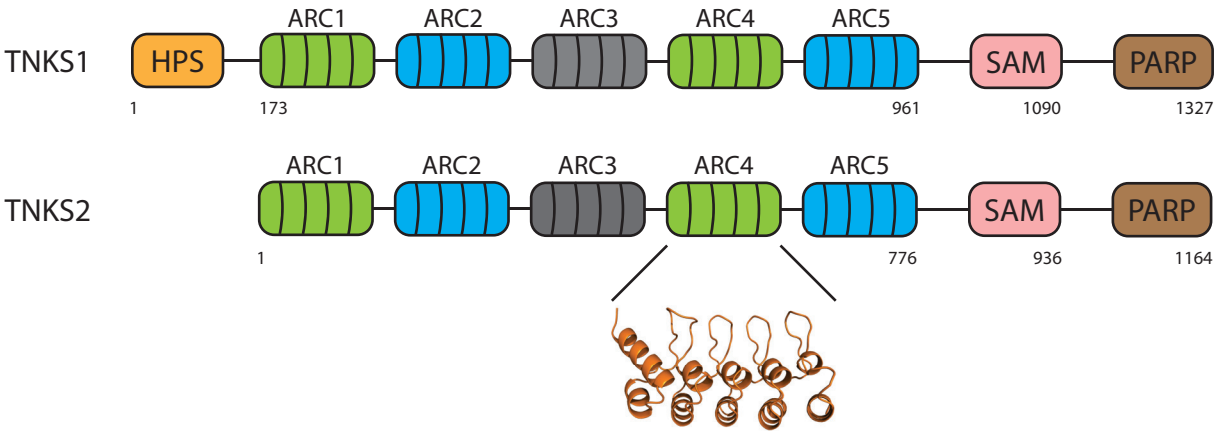


Figure 1.4: Domain structure of Tankyrase 1 (TNKS1) and Tankyrase 2 (TNKS2). TNKS1 contains a unique HPS domain. Both proteins contain an ANK domain consisting of five so-called ankyrin repeat clusters (ARCs), a sterile alpha motif (SAM) domain and a catalytic PARP domain. All ARCs, excluding ARC3 (grey) are able to bind substrate proteins. Approximate domain boundaries are highlighted for both Tankyrase proteins. Not all start and ends of boundaries have been included for each domain for clarity. Schematic of the x-ray crystal structure of TNKS2 ARC4 (PDB ID: 5BXO) is also shown.

Within the human genome, there are two TNKS proteins, TNKS1 and TNKS2 (Figure 1.4). Both proteins share high sequence homology and structural domains. TNKS1 is unique as it contains a Histidine Proline Serine-rich (HPS) region at its N-terminus. Both proteins share an ANK domain comprising five ankyrin repeat clusters (ARCs), a SAM domain involved in self-oligomerisation and a PARP domain at the C-terminus. Each of these shared domains has high homology, with 83%, 74% and 94% sequence identity

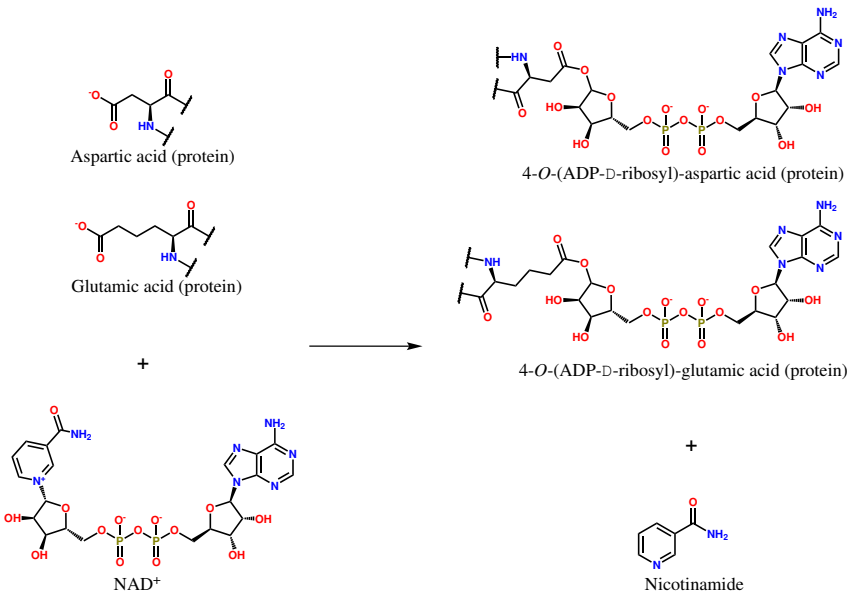


Figure 1.5: Reaction schematic for enzymatic mono-ADP ribosylation of substrate proteins. Mono-ADP ribosylation occurs on substrate aspartic and/ or glutamic acid residues. Mono-ADP-ribosylated are further elongated by the PARP domain.

between the ANK, SAM and PARP domains, respectively [125]. The role of the HPS domain remains unknown to date, although it is expected to be intrinsically disordered [126]. The SAM domain is known to modulate TNKS oligomerisation [127, 128], but until recently its exact function was still largely ambiguous. More recent work by the Guettler laboratory defined the SAM domain as crucial in the formation of localised signalling puncta [129]. TNKS1 and TNKS2, also known as PARP-5a and PARP-5b, respectively PARylate substrate proteins through the reaction depicted in Figure 1.5.

1.4.2 Tankyrase cellular function and signalling

Tankyrase substrates are recruited to ARCs 1, 2, 4, and 5 through their short linear motifs (SLiMs). Early work from Sbodio and Chi described a minimal hexapeptide motif of RxxPDG (where x represents any amino acid) that was derived from Nuclear Mitotic apparatus protein 1 (NuMA) that was reported to bind the ARC domains [130]. An optimised motif was reported by Guettler *et al.*., which comprises an eight-residue peptide (REAGDGEE) that bound the ARC domains with approximately 10-fold higher affinity [131]. All peptides bind in a largely conserved surface of the ARCs, in which the canonical contacts are defined by the guanidino group of Arg1 binding the 'arginine cradle site'. In addition, Gly6 is wedged between the sidechains of Tyr or Phe that form $\pi - \pi$ stacking interactions (Figure 1.6).

PARylated substrates of TNKS are recognised by RING finger protein 146 (RNF146),

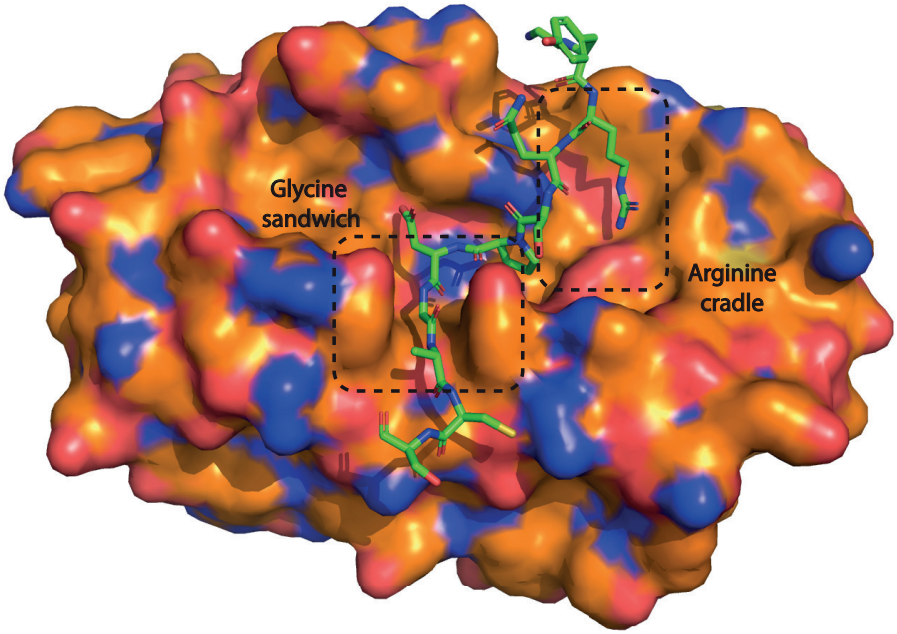


Figure 1.6: Canonical binding of the IRAP peptide to TNKS1 ARC2, described by Arg1 sitting in the arginine cradle site and Gly6 situated between Tyr379 and Tyr412. Arginine cradle and glycine sandwich are highlighted in the dashed box. PDB ID: 5JHQ [126].

an E3 ubiquitin ligase that specifically binds the iso-ADP-ribose moiety of the poly-ADP-ribose chain and ubiquitinates the substrate, leading to its subsequent degradation by the 26S proteasome [132]. Interestingly, TNKS regulate themselves by autoPARylation resulting in their ubiquitination by RNF146 and downstream degradation [122, 133]. More recently it was discovered that RNF146 binds to Tankyrase at the substrate-binding sites via up to four non-canonical SLiMs [134]. These sequences contain either one or two extra amino acids between positions 1 and 4 of the canonical binding peptide. Currently, it is not understood whether all four sequences are used for binding TNKS or if the interaction is of biological relevance. However, if such interactions do occur, the presence of a ternary complex containing TNKS, RNF146 and TNKS substrates may provide an explanation for PARP-independent substrate degradation.

Tankyrases have roles in several cellular pathways including the regulation of glucose uptake [121], telomere maintenance [130, 133], DNA repair [133, 135], and cancer (reviewed in [112, 136], and references therein, see Figure 1.7). The interaction between TNKS and Axin1 (the most studied of the Axin isoforms) within the WNT signalling pathway has garnered the most attention in the oncology field. This is due to its pivotal role in the formation of the β-catenin destruction complex, where Axin1 is a limiting factor in the destruction complex formation. Axin1 binds to TNKS through two non-

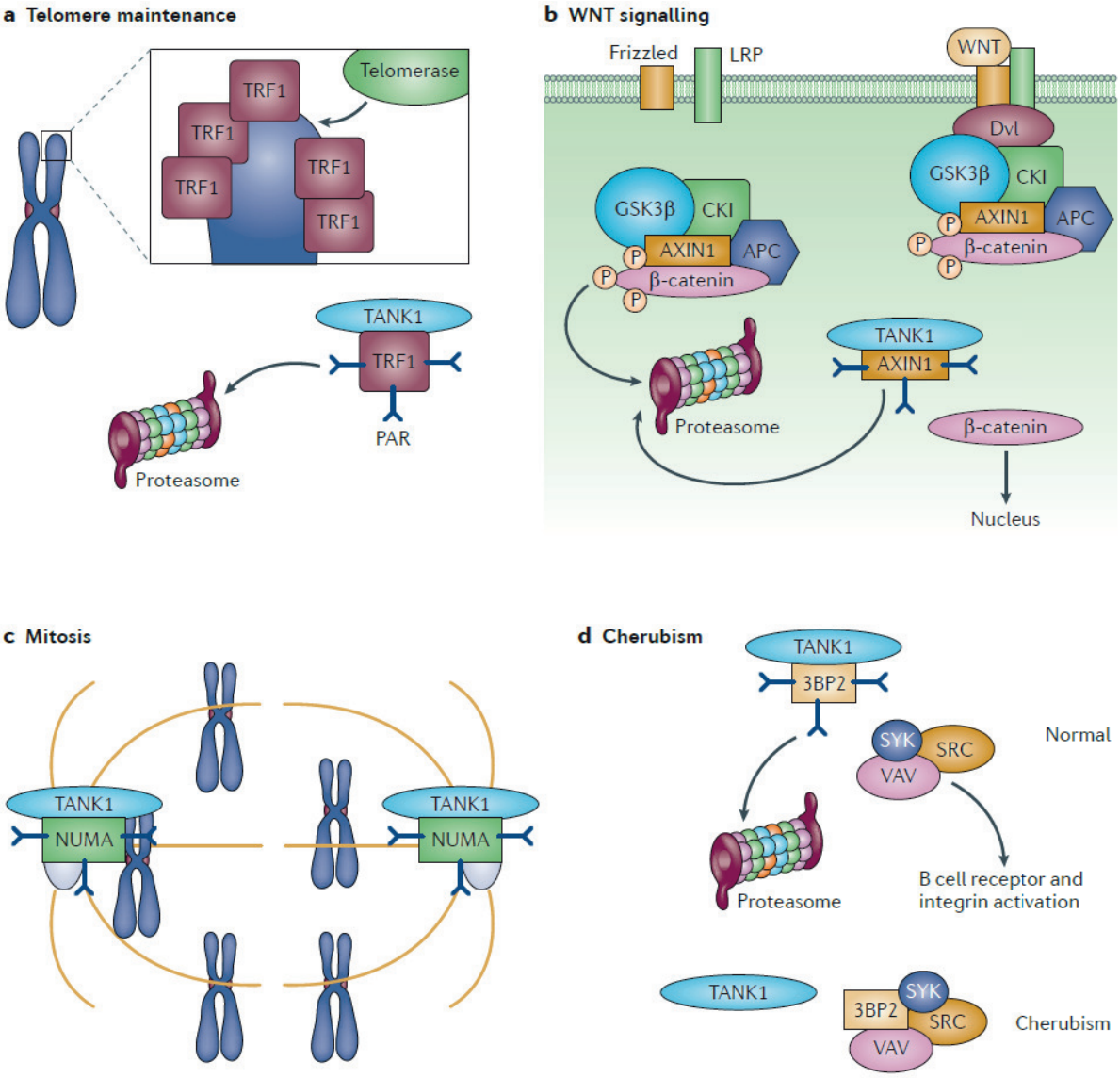


Figure 1.7: Overview of normal cellular processes and disease pathologies associated with TNKS function. (a) Tankyrase 1 (TANK1) PARylates telomeric repeat binding factor 1 (TRF1) resulting in its degradation by the 26S proteasome. Telomerase can then access and extend the telomere. (b) TANK1/2 PARylate Axin1 resulting in its ubiquitination and degradation by the 26S proteasome, allowing β-catenin to translocate to the nucleus and activate the transciption of downstream oncogenes. (c) TANK1 PARylates NuMA among other mitotic proteins. (d) TANK1 PARylates 3BP2 targeting it for proteasomal degradation resulting in normal downstream function. Mutations in 3BP2 blocking PARylation result in a lack of degradation and lead to cherubism. Figure adapted from Riffel *et al.*. [112].

identical, canonical SLiMs in its N-terminal region (residues 1-80), as seen in the X-ray crystal structures of mTNKS1 ARC2-3 and mAxin1 1-80 [118].

Small-angle X-ray scattering (SAXS) in and X-ray crystallography data of provided

further structural insights between the ARC domains and Axin [126]. In the context of the full-length ANK domain (ARCs 1-5), Axin1 binds preferentially to ARC2 and ARC5 simultaneously. It was demonstrated that shortening the linker regions between the two canonical tankyrase-binding peptides of Axin1 can induce binding between adjacent ARC1 and ARC2 or ARC4 and ARC5 domains. These findings suggest the length of the linker between two SLiMs of a single TNKS substrate could be functionally relevant. Lastly, conformational changes in the TNKS ANK domain were observed as a result of binding Axin1, though it is not yet understood if this is relevant to its PARP activity [137].

Tankyrases, in addition to other PARP domain-containing proteins, have been the subject of multiple drug-discovery campaigns in both academia and the pharmaceutical industry. Inhibition of the catalytic activity of PARP domains is a highly attractive therapeutic strategy such as that of PARP1 in BRCA1/2 malignancies [138, 139]. PARP1 is highly involved with the sensing of DNA single-stranded breaks, which is a hallmark of genome instability that is often associated with cancers [140]. Inhibition of PARP1 has been a major focus for pharmaceutical companies and has resulted in a large number of molecules being produced preferential for PARP1, such as Olaparib (AZD-2281, AstraZeneca) and Veliparib (ABT-888, AbbVie) among many others listed in references [112, 141]. Most TNKS PARP inhibitors target either the nicotinamide-binding site (XAV939 [120, 142]) or the adenosine-binding site (IWR-1 [143], G007-LK [144], WIKI4 [145]), and some inhibitors are able to target both (PJ34 [142], Long quinazolinone [146]). The molecular mechanism or therapeutic strategy between targeting different PARPs is very different owing to the varying functions of individual PARP domain-containing proteins. However, the substrate binding domains have often been neglected. This is largely due to the binding interface being large, flat and extended, classifying it as 'undruggable' by small molecule therapeutics. To date, targeting the ARC domains has been explored by fragment-screening [147] and macrocyclised peptide approaches [102], of which the latter resulted in the most potent inhibitors.

1.5 Anaphase Promoting Complex/Cyclosome

The discovery of the Anaphase Promoting Complex/Cyclosome (APC/C) was reported independently by Sudakin *et al.*. and King *et al.*. in 1995 [148, 149]. They were the first to identify the large multi-protein complex of approximately 1.5 MDa as an E3 ubiquitin ligase that ubiquitinates cyclin B in mitosis, thereby mediating its proteasomal degradation. Since its discovery, significant technological advances have enabled researchers in the field to gain structural insights that underpin APC/C function as a master regulator of the cell cycle.

1.5.1 Structure of the APC/C

The groups of David Barford and Brenda Schulman have focused on gaining a structural and mechanistic understanding of the APC/C using X-ray crystallography, negative-stain electron microscopy and cryogenic electron microscopy (Cryo-EM) [28, 150–154]. The core APC/C comprises a total of 19 subunits [155], of which 14 are unique and five are present as dimers (Table 1.2). An illustration of the molecular architecture of the APC/C^{Cdh1} is shown in Figure 1.8; in general the APC/C is subdivided into the Platform, the TPR (tetratricopeptide repeat) lobe, which enclose the catalytic and substrate binding subunits at its centre. The co-activator subunits (Cdc20/Cdh1), which interact with subunits in the TPR lobe, are responsible for recruiting substrates for ubiquitination by the E2 enzymes. Ube2C adds the initial ubiquitin monomer [156], while Ube2S is primarily responsible for ubiquitin-chain elongation via K11-linkages [157]. Ubiquitin chain branching is also catalysed by Ube2C [155].

Table 1.2: Subunits comprising the APC/C. Abbreviations: IR-tail - Isoleucine-Arginine tail motif, MCC - Mitotic Checkpoint Complex, PC - Proteasome/cyclosome repeat, TPR - Tetratricopeptide repeat, RING - Really Interesting New Gene. Table adapted from [155].

Subunit	Mass (kDa)	N	Function	Location	Domains
Apc1	216.4	1	Scaffolding	Platform	WD40, mid-helix, PC
Apc4	92.1	1	Scaffolding	Platform	WD40, four helix bundle
Apc5	85.1	1	Scaffolding	Platform	N-terminal helical domain, 13 TPR
Apc3	91.9	2	Scaffolding	TPR lobe	14 TPR
Apc6	71.7	2	Scaffolding	TPR lobe	14 TPR
Apc7	66.9	2	Scaffolding	TPR lobe	14 TPR
Apc8	68.8	2	Scaffolding	TPR lobe	14 TPR
Apc12	9.8	2	TPR-accessory	TPR lobe	α-helix
Apc13	8.5	1	TPR-accessory	TPR lobe	α-helix
Apc16	11.7	1	TPR-accessory	TPR lobe	α-helix
Apc2	93.8	1	Catalytic	Catalytic module	Cullin
Apc11	9.8	1	Catalytic	Catalytic module	RING
Apc10	21.2	1	D-box recognition	Substrate binding	IR-tail
Apc15	14.3	1	MCC interaction	Platform	α-helix
APC/C	1167.4	19	-	-	-

Both co-activator proteins (Cdc20 and Cdh1) bind to the same site on the APC/C, whereby a region of their respective N-terminal intrinsically-disordered regions (IDRs) associates with both Apc1 and Apc8, becoming partially structured [150, 151, 159, 160].

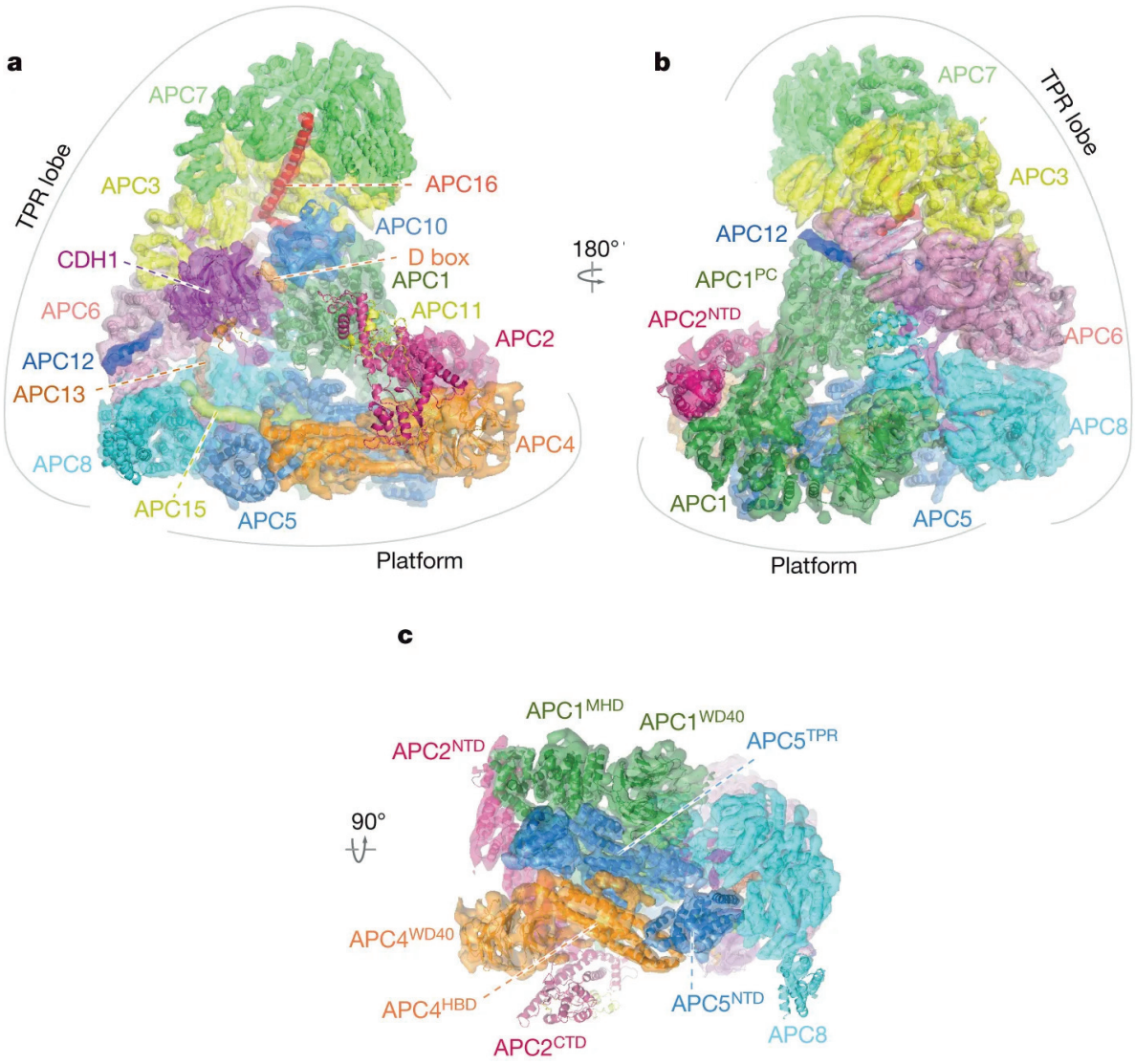
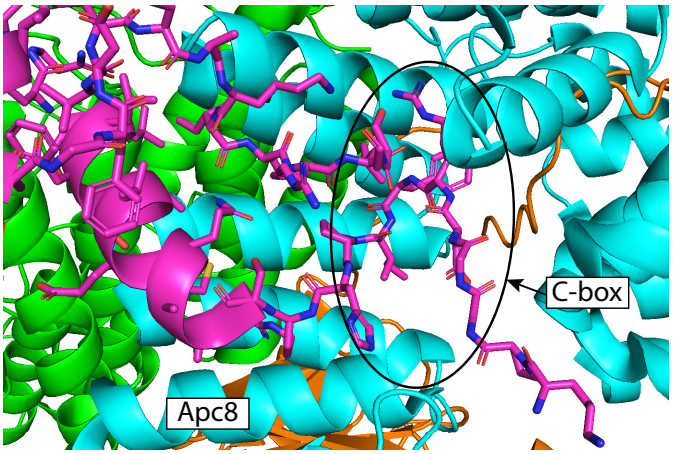
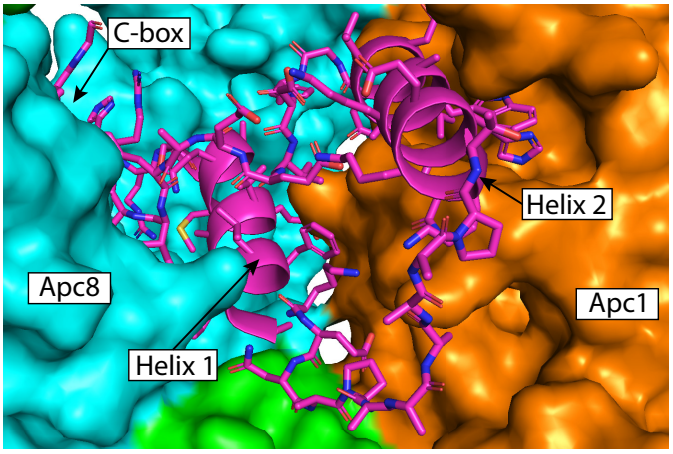


Figure 1.8: Cryo-EM model of the APC/C^{Cdh1}-Hsl1 ternary complex at 7.4 Å resolution. (a) Front, (b) back and (c) Bottom views of the APC/C. Substrate binding and ubiquitination occurs in the space framed by the Apc2, Apc10 and Cdh1/Cdc20 triad. Figure adapted from [158].

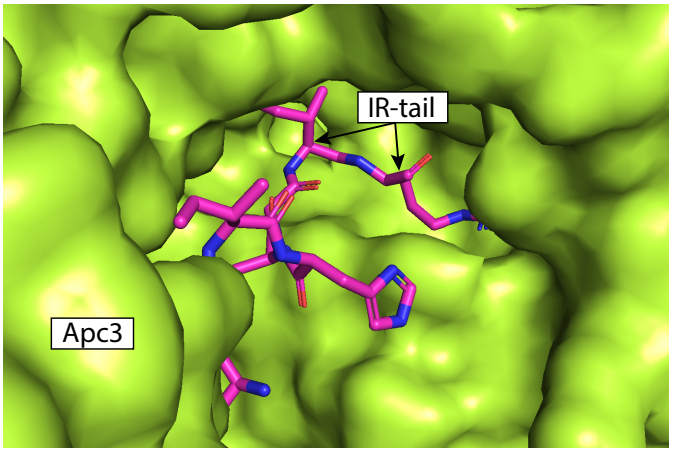
The so-called C-box motif buries into a pocket in Apc8 in a loop conformation (Figure 1.9a). Two other regions of the N-terminal IDR adopt α-helical structures, Helix 1 and Helix 2, that bind the interface between Apc1 and Apc8, and the surface of Apc1, respectively (Figure 1.9b). In addition to the N-terminal domain, both Cdc20 and Cdh1 share a conserved isoleucine-arginine dipeptide motif (IR-tail), which binds to Apc3 in a region that is structurally homologous to the C-box binding site of Apc8 (Figure 1.9c) [150, 151, 161, 162].



(a) Cdc20 C-box motif



(b) Cdc20 Helix interactions



(c) Cdc20 IR-tail

Figure 1.9: Co-activators bind to the APC/C using conserved motifs contained in their N- and C-terminal regions. (a) The Cdc20 C-box motifs forms a tight loop configuration burying R78 and Y79 into a groove formed by Apc8. (b) Cdc20 α-helices (magenta, cartoon and stick representations) bind to sites formed by Apc1 (orange, surface) and the interface between Apc1 and Apc8 (cyan, surface). (c) The Cdc20 IR-tail binds the structurally homologous site of Apc3 (lime green, surface). Images were generated from PDB:5LCW [151].

1.5.2 Cell cycle regulation by the APC/C

The APC/C temporally regulates various cell cycle proteins by ubiquitination leading to their subsequent proteasome-mediated degradation (Figure 1.10) [163]. The APC/C is able to exert its E3 ligase activity via its co-activator proteins that recruit substrates through three principal peptide sequences, known and hereby referred to as degrons: the Destruction box (D-box), Lysine-Glutamate-Asparagine-box (KEN-box) and ABBA motif [164–167], which are further discussed in Chapter 5. Identification of APC/C substrates and their degrons is a very active field of research, and their numbers are steadily increasing. A repository containing all known APC/C substrates has been collated by Davey and Morgan [168, 169].

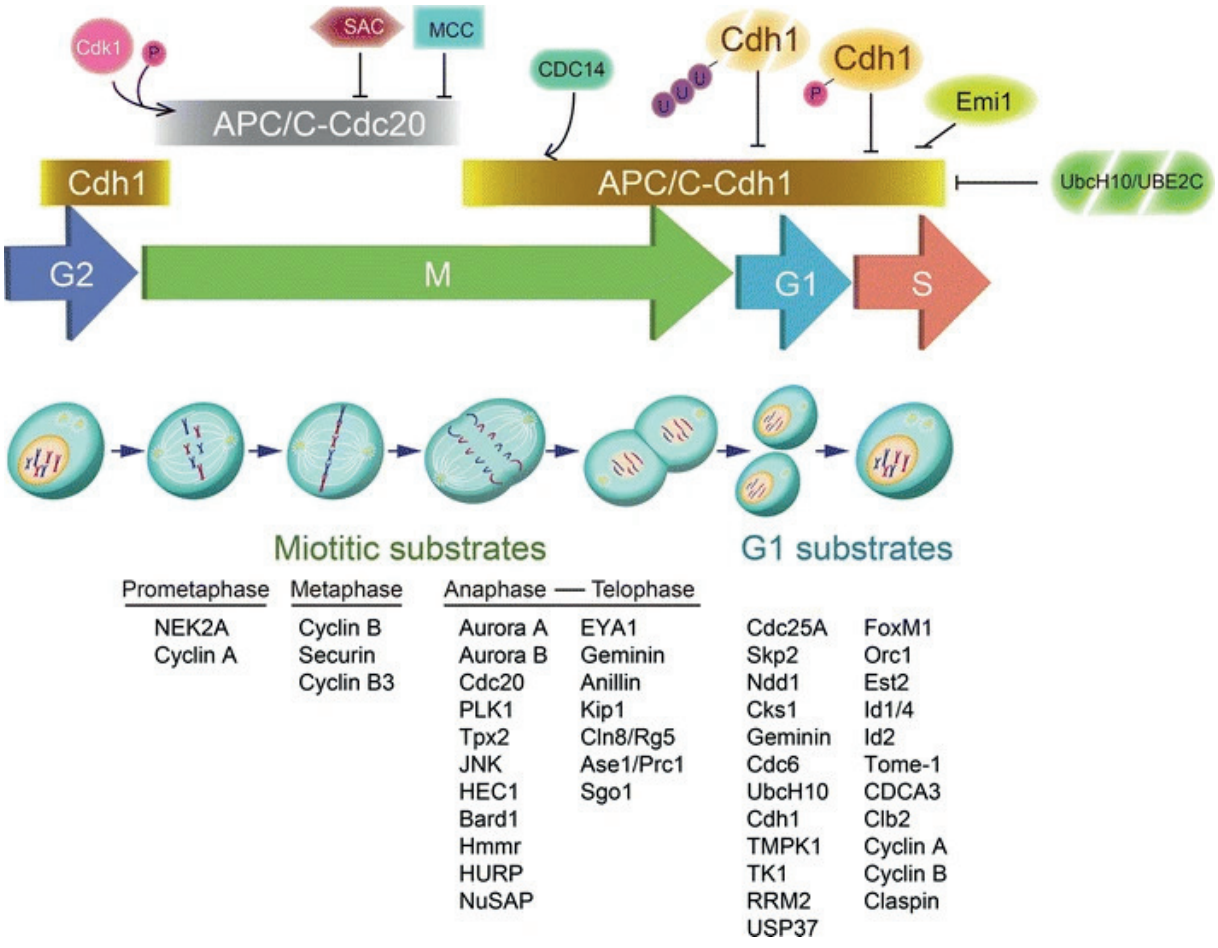


Figure 1.10: The APC/C regulates the cell cycle through ubiquitination and proteasome-mediated degradation of cell cycle proteins. Figure adapted from Zhou *et al.*. [163].

Prior to the onset of the mitosis, Cdc20 is phosphorylated by Cyclin-dependant kinase 1 (Cdk1), which initiates the association of Cdc20 with the APC/C and thereby activation

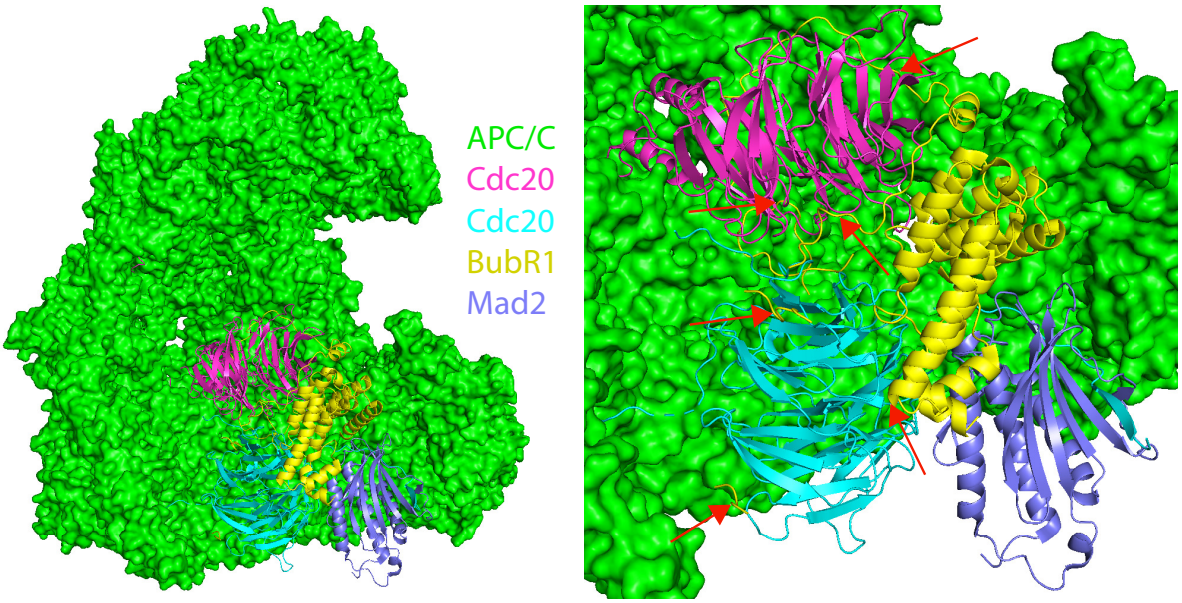


Figure 1.11: The structural basis for MCC-mediated inhibition of the APC/C^{Cdc20}. (a) Overall architecture of the MCC bound to the APC/C^{Cdc20}. (b) The BubR1 subunit of the MCC binds to Cdc20 molecules within both the MCC and the APC/C using pseudo-degron peptides (indicated by red arrows). Figures produced using PDB:5LCW [151].

of the APC/C^{Cdc20} [170, 171]. The APC/C^{Cdc20} remains active, degrading Nek2A [172] and Cyclin A [154], until the spindle is correctly aligned and kinetochores are correctly attached to sister chromatids in pro-metaphase. The APC/C^{Cdc20} is tightly regulated until the spindle assembly checkpoint (SAC) is satisfied, after which degradation of Cyclin B1 and securin occurs allowing the cell to proceed from metaphase into anaphase [163]. After the SAC arrest, the APC/C^{Cdc20} is inhibited by the Mitotic Checkpoint Complex (MCC), which contains BubR1, Cdc20 and Mad2A [160, 173]. Alfieri and co-workers determined the structure of the APC/C engaged with the MCC (APC/C^{MCC}), providing evidence for the molecular mechanism by which the MCC inhibits the APC/C^{Cdc20} [151]. BubR1 interacts with the APC/C-bound Cdc20 while simultaneously engaging MCC-bound Cdc20. This interaction across both complexes is mediated by pseudo-degron motifs in BubR1 that encompass two D-box, two KEN-box and two ABBA motif peptides (Figure 1.11). Cdc20 is auto-ubiquitinated in an Apc15-mediated interaction, and replaced by de-phosphorylated Cdh1 in the remainder of mitosis [174–176]. The APC/C^{Cdh1} maintains low levels of Cdc20 until S-phase by targeting it for ubiquitination to prevent premature mitotic entry [177].

1.5.3 APC/C in oncology

Although temporal regulation of Cdc20- and Cdh1-bound APC/C is tightly regulated in healthy cells, the opposite occurs in the vast majority of cancer cells. APC/C^{Cdh1} provides tumour suppressive functions, e.g. by inhibiting premature mitotic entry, and hence Cdh1 is often found to be down-regulated in cancers [163]. However, Cdc20 drives mitotic entry and therefore is often found to be up-regulated at the mRNA and protein levels of patients with a wide variety of cancers, which include those listed in Figure 1.12 and several more [178–185]. Such widespread upregulation of the Cdc20 across multiple cancer tissues implies necessity of its function to cancer progression. Indeed its overexpression has already been implicated in poor prognosis for patients with colorectal cancer [181], breast cancers [186], cutaneous squamous cell carcinomas [187] and various solid tumours [182]. In addition, many of these reports indicate that Cdc20 may prove to be a valuable biomarker for the diagnosis of these diseases. While most studies have focused on the relative levels of Cdc20, none have attempted to identify whether the CDC20 gene is particularly prone to mutation that can cause loss or gain of function. To investigate this, clinical datasets from The Cancer Genome Atlas (TGCA, [188], the Catalogue Of Somatic Cancer Mutations in Cancer (COSMIC, [189]) and the International Cancer Genome Consortium (ICGC, [190] were analysed using BioMuta version 4.0 [191] for the search query of the CDC20 and TP53 genes, focusing of hotspot or frequently mutated residues (Figure 1.13.

Of particular importance to note, the TP53 is commonly known to be mutated in many cancers and this can be clearly observed in Figure 1.13b [192]. In stark contrast, paying particular attention to the scale of y-axis, i.e. the frequency of variations found in cancer patients, the CDC20 gene is far less commonly mutated when comparing the same datasets (Figure 1.13a). It is difficult to evaluate whether the frequency of mutations seen in CDC20 are truly relevant to the the progression of oncology. However, it appears that R445 has a tendency to mutate to either glutamine or tryptophan in these samples. R445 lies within the WD40 domain, and so it was hypothesised whether this mutation might be directly linked with degron binding. All the residues mapped in Figure 1.13a were then subsequently mapped onto the crystal structure of human Cdc20 bound to a Bub1B-derived KEN-box peptide shown in Figure 1.14 [193]. R445 maps directly to a intermolecular hydrogen bond from the Arginine 445 side chain to the conserved Glutamic acid carbonyl and Aspargine carboxamide group on the side chain. It is possible that mutation of R445 to Q or W could be detrimental, but it could also be beneficial. However, given that there is no further literature to suggest that the R445Q/W mutations is functionally relevant to the activity of the APC/C^{Cdc20} in its ubiquitination of substrates

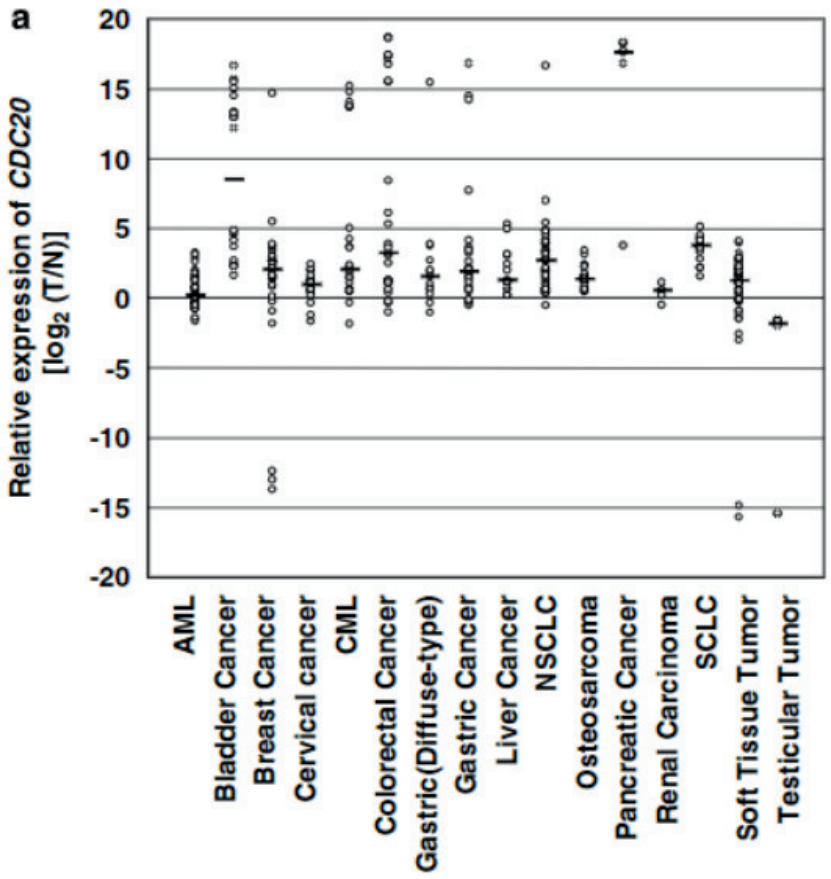
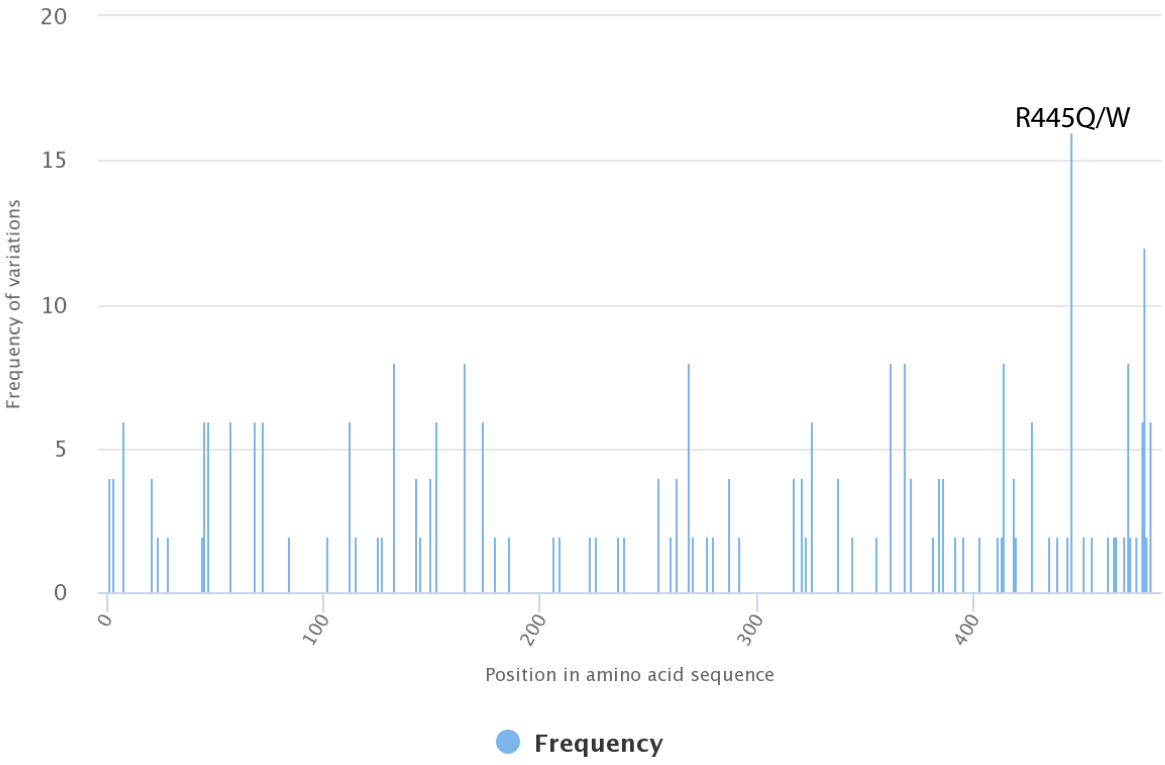


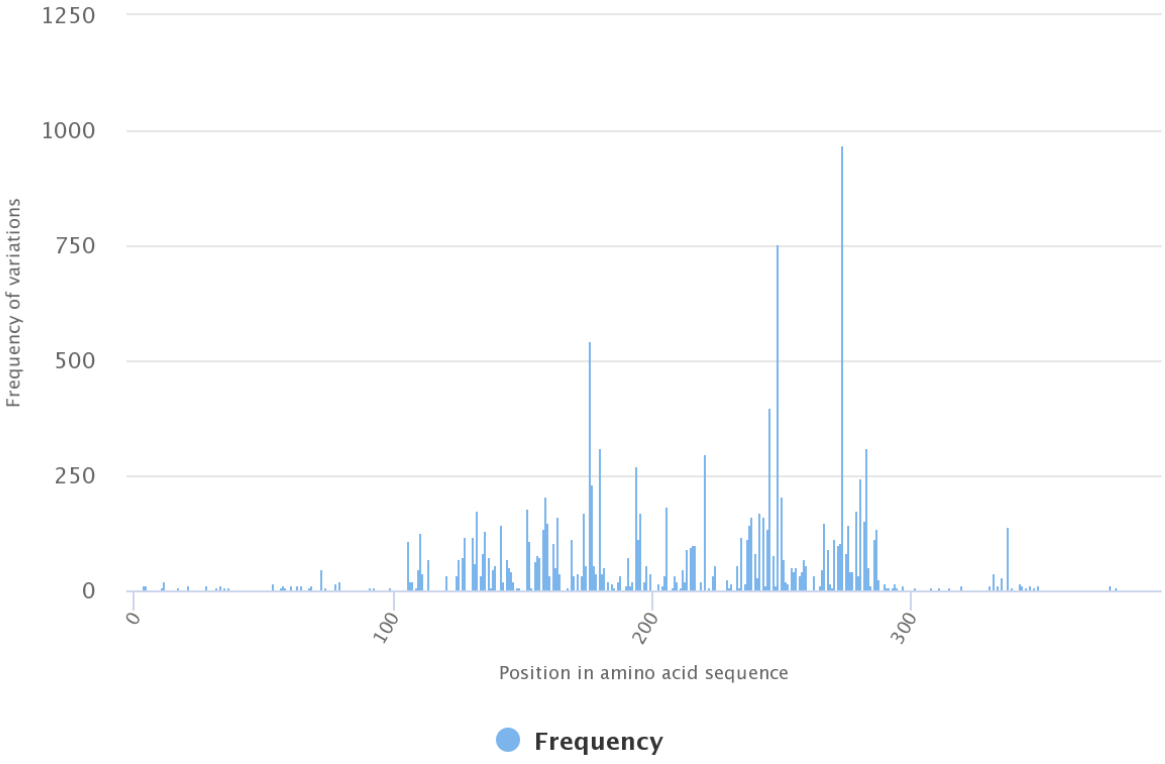
Figure 1.12: Overexpression of Cdc20 from cDNA microarray analysis. Levels are relative between tumourigenic tissue (T) and non-tumourigenic tissue (N) across the various tissue types. Figure adapted from Kidoro et al. [178].

mediated by the KEN-box degron, further speculation is severely error prone. The lack of hotspot mutations in CDC20, in comparison to the TP53 gene, is perhaps unsurprising. One must consider that p53 acts as a tumour suppressor protein and so mutation of this function can result in oncogenesis. Conversely, wild-type Cdc20 is itself a driver of mitosis and thus simple overexpression is sufficient to drive oncogenesis. In addition, it makes very little sense to knockout any degron binding sites as the ubiquitination and proteasomal degradation of substrates is a crucial factor in the progression of the cell cycle.

Despite its pivotal role in many cancers, very few inhibitors against the APC/C^{Cdc20} exist, none of which are in clinical trials or are approved drugs at present. To date, only two compounds, Apcin and TAME (and its pro-drug analog, proTAME), were identified as competitive inhibitors against the APC/C^{Cdc20} (Figure 1.15). They were originally discovered in a high-throughput screen comprising 109,113 molecules, that monitored the drug-induced stabilisation of a Cyclin B1 (2-97)-luciferase reporter in Xenopus extracts as a readout of the inhibition of mitotic exit [197]. While the original article describing the discovery of these compounds include Apcin and TAME (referred to as the Class



(a) CDC20



(b) TP53

Figure 1.13: Frequency plot of mutations occuring in the (a) CDC20 (Uniprot ID: Q12834) and (b) TP53 (Uniprot ID: P04637) genes from clinical datasets of human cancers [188–190]. Figures generated using BioMuta version 4.0 [191].

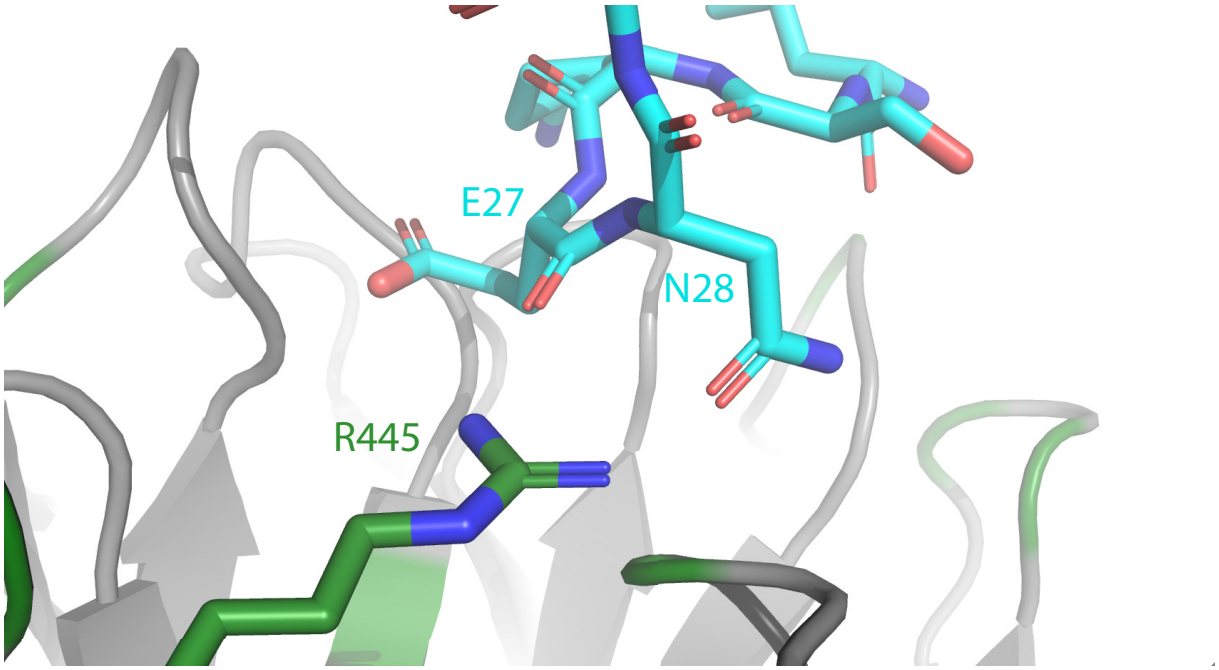


Figure 1.14: Somatic mutation of R445Q/W may result in altered binding to substrate proteins containing KEN-box degron motifs. Cdc20 in grey (mutations are highlighted in green according to Figure 1.13a, and the Bub1B KEN-box peptide bound to Cdc20. E27 carbonyl and N38 carboxamide form intermolecular H-bonds with R445. PDB: 4GGD, [193].

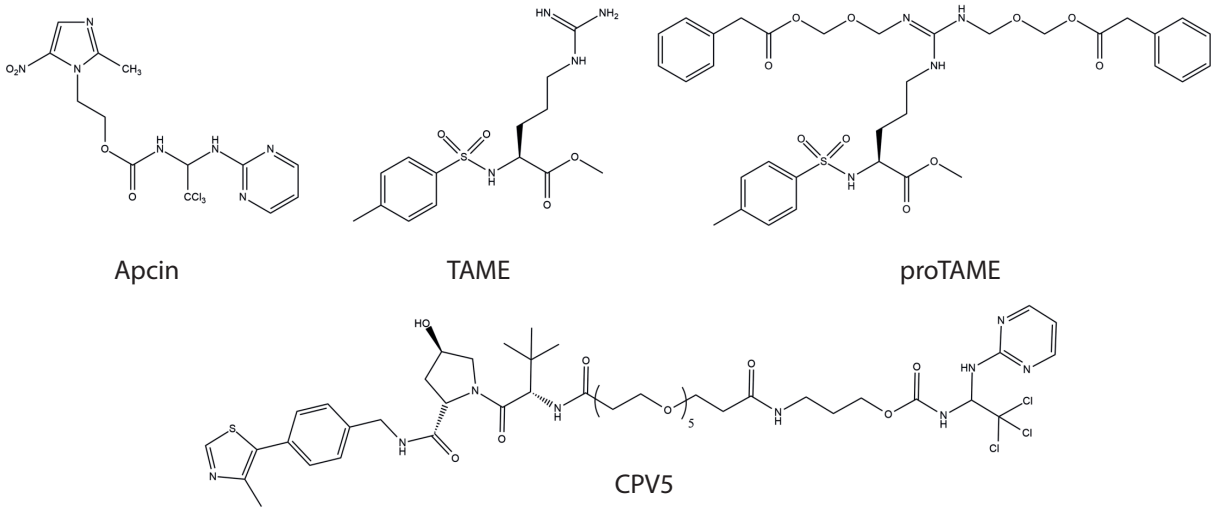


Figure 1.15: Structures of current validated APC/C^{Cdc20} inhibitors. Apcin competes for the D-box binding site of Cdc20 and TAME/proTAME competes for the C-box and IR-tail binding sites on Apc8 and Apc3[28, 194, 195]. The PROTAC molecule CPV5 has also shown efficacy *in vivo* to reduce tumour size in mouse xenograft models [196].

*II*A compounds), the authors focus is rather towards the Class *II*B compounds that they describe, which were not designated as Cyclin B1 ubiquitination inhibitors. The mechanisms of action for Apcin and TAME remained unknown for several years following their

discovery. Further investigation revealed TAME to compete for co-activator (Cdc20 and Cdh1) binding to the APC/C [194, 198]. TAME in it's active form is not cell permeable and it therefore the proTAME derivative was made containing N,N'-bis(acyloxymethyl carbamate) groups, which can be cleaved by intracellular esterases [194]. However, since these broad specificity esterases also exist in the serum, proTAME will be modified before it reaches the intracellular space, posing a major problem for the molecule. The second molecule, Apcin was shown to be a competitive inhibitor for the D-box degron-binding site of Cdc20 (with limited binding to Cdh1) [195]. The functional activity of Apcin is however not so straight forward and its effects can produce mitotic arrest, but also increase the rate of mitosis under high SAC activity [199]. However, the use of both Apcin and proTAME in combination, have a synergistic effect on the inhibition of mitotic exit due to their alternative inhibitory mechanisms [195]. One recent study developed a proteolysis targeting chimera (PROTAC) molecule, CP5V, by coupling Apcin to a Von-Hippel Lindau (VHL) ligand using a five-membered polyethylene glycol $(PEG)_5$ linker. In response to CP5V treatment, degradation of Cdc20 was observed in multiple cancer cells lines, and tumour sizes were reduced in 4T1 xenograft mouse models [196]. Even though CP5V is provides evidence that degradation of Cdc20 is promising as a mode of therapeutic intervention, the failure rate of compounds reaching and succeeding in clinical trials is high. Due to the poor efficacy of Apcin as a monotherapeutic agent, issues with cell permeability of TAME/proTAME, and the lack of ligand diversity available for use in PROTACs such as CPV5, the development of novel molecules is desperately needed to pave the way towards new therapeutic agents that are required across a large number of cancers where Cdc20 activity is high.

1.6 Project aims and objectives

The inhibition of protein-protein interactions (PPIs) in biological systems has presented a challenge to small molecule drug campaigns. In the context of cancer, targeting these interactions has been largely unmet and requires alternative strategies. Two such targets are Tankyrase and the APC/C^{Cdc20}, and a major goal of my doctoral studies was to design peptides that could function as inhibitors of these proteins.

The work encompassing the development of Tankyrase inhibitory peptides was built upon previous work performed by Dr. Wenshu Xu in the Itzhaki laboratory. Although peptides of sub-micromolar affinities had already been obtained and cell permeability had been achieved, their cellular efficacies required improvement. Therefore, the main aim of the Tankyrase sub-project was to develop peptides with higher affinities. To achieve this aim, both the consensus sequence and the macrocycle positions required further optimisation (Chapter 3).

The main focus of the PhD, was to target the APC/C^{Cdc20}'s oncogenic activity. The aims of the project were to:

1. Design and synthesise peptides to bind the E3 ubiquitin ligase APC/C^{Cdc20},
2. Characterise the binding of the peptides to Cdc20 using biophysical methods,
3. Investigate their inhibitory function in a biological context.

Structure-based drug design and consensus analyses in combination with various modelling approaches formed the basis of the design process (Chapter 5). Given the vast chemical space that can be explored within peptide chemistry, it was logical to also employ macrocyclisation and the incorporation of non-natural amino acids. In order to characterise Cdc20-peptide interactions and thereby understand the structure-activity relationship, it was necessary to express and purify recombinant Cdc20 (Chapter 4) and establish biophysical assays that could probe the interactions (Chapter 6). The aim was to test the inhibitory function of lead peptides using in vitro ubiquitination assays (in collaboration with David Barford, MRC-LMB) and cellular assays to examine the inhibition of mitosis (in collaboration with Catherine Lindon, Department of Pharmacology, University of Cambridge). However, due to time constraints these experiments were not conducted prior to submission of the thesis.

Chapter 2

General materials and methods

2.1 Buffers and reagents

All chemicals were purchased from Sigma-Aldrich, Thermo Fisher Scientific or Merck, unless otherwise stated. Oligonucleotides for cloning were purchased from either Sigma-Aldrich or Integrated DNA technologies (IDT). All molecular biology enzymes were purchased from New England Biolabs (NEB) or Thermo Fisher Scientific. Molecular grade agarose was purchased from Bioline. 2x Yeast-Tryptone (2xYT) powder was purchased from Formedium. All chromatography columns were purchased from GE Healthcare. Polyethersulfone (PES) centrifugal concentrators were purchased from Sartorius Stedim. All primary antibodies were puchased from Cell Signalling Technologies (CST) and secondary antibodies from DAKO. General buffers and stock solutions are listed below. All solutions were prepared using MilliQ H_2O (18.2 MΩ/cm), unless otherwise stated. Vectors and *E. coli* strains used are listed in Tables 2.1 and 2.2, respectively.

- Ampicillin: 100 µg/mL, sterile filtered.
- Kanamycin: 50 µg/mL, sterile filtered.
- Chloramphenicol: 34 µg/mL in ethanol.
- Ethylenediaminetetraacetic acid (EDTA): 0.5 M, pH 8.0 (adjusted with NaOH), sterile filtered.
- Isopropyl-β-D-thiogalactoside (IPTG): 1 M, sterile filtered.
- Dithiothreitol (DTT): 1 M, sterile filtered.
- Tris(2-carboxyethyl)phosphine (TCEP): 0.5 M, pH 7.0 (adjusted with NaOH), sterile filtered.
- P2 buffer: 200 mM NaOH and 1% (w/v) sodium dodecyl sulfate (SDS)

- N3 buffer: 4.2 M guanidinium hydrochloride and 0.9 M potassium acetate (KOAc), pH 4.8 (adjusted with glacial acetic acid).
- DNA wash buffer: 10 mM Tris-HCl, 80 % (v/v) ethanol, pH 7.5
- Lysogeny broth (LB): 25 g/L, autoclaved 121°C, 15 minutes.
- 2xYT: 31 g/L, autoclaved 121°C, 15 minutes.
- LB-Agar: 35 g/L, autoclaved 121°C, 15 minutes. Appropriate antibiotics added when cooled below 50°C, if required.
- Transformation buffer I (TfbI): 30 mM KOAc, 100 mM RbCl, 10 mM $CaCl_2$, 50 mM $MnCl_2$, 3 mM $[Co(NH_3)_6]Cl_3$ and 15% (v/v) glycerol, pH 5.8 (adjusted with glacial acetic acid). Sterile filtered.
- Transformation buffer II (TfbII): 10 mM 3-(N-morpholino)propanesulfonic acid (MOPS), 10 mM RbCl, 75 mM $CaCl_2$ and 15% (v/v) glycerol, pH 6.5 (adjusted with KOH). Sterile filtered.
- Tris-acetate-EDTA (TAE) buffer: 40 mM Tris base, 1 mM EDTA, (adjusted with 5.71 mL glacial acetic acid). Sterile filtered.
- Separating gel buffer: 1.5 M Tris-HCl and 4% (w/v) SDS, pH 8.8.
- Stacking gel buffer: 0.5 M Tris-HCl and 4% (w/v) SDS, pH 6.8.
- 2× SDS PAGE loading buffer: 100 mM Tris-HCl, 4% (v/v) SDS, 0.2% (w/v) bromophenol blue, 30% (v/v) glycerol and 715 mM β-mercaptoethanol, pH 6.8.
- Tris-Glycine-SDS (TGS) buffer: 25 mM Tris base, 192 mM glycine and 1% (w/v) SDS, pH 8.8.
- Transfer buffer: 25 mM Tris base, 192 mM glycine, 20% (v/v) MeOH, pH 8.4
- Phosphate-buffered saline (PBS): 8 mM Na_2HPO_4, 2 mM KH_2PO_4, 137 mM NaCl and 2.7 mM KCl, pH 7.4.
- Phosphate-buffered saline-Tween20 (PBST): 8 mM Na_2HPO_4, 2 mM KH_2PO_4, 137 mM NaCl, 2.7 mM KCl and 0.1% (v/v) Tween 20, pH 7.4.
- Western blot blocking buffer: 5% (w/v) Marvel non-fat dry milk powder in PBST.
- Coomassie brilliant blue G-250: 80 mg/L, add 3 mL concentrated HCl.

Vector	Resistance	Description	Source
pRSET	Ampicillin	High-copy number vector encoding for N-terminal His_6-tagged recombinant proteins, with a thrombin cleavage site between His_6-tag and recombinant protein. Expression controlled by the T7 promoter.	Itzhaki lab
pRSET-TEV	Ampicillin	As above, except thrombin cleavage site is replaced by a tobacco etch virus (TEV) protease cleavage site.	R.S. Eapen
pGST	Ampicillin	As pRSET, except His_6-tag is replaced by a glutathione S-transferase tag.	Itzhaki lab
pOWEN	Ampicillin	Derivative of pOPIN F vector. High-copy number vector encoding for N-terminal His_6-tagged recombinant proteins with a thrombin cleavage site between His_6-tag and recombinant protein. Expression of proteins in *E. coli* (T7), baculovirus (p10) and mammalian systems (CAG).	Dr. O.D. Burbidge
pTriExMOD	Ampicillin	Modified pOPIN E vector. High-copy number vector. CAG promoter replaced by the human cytomegalovirus (CMV) immediate early enhancer and CMV promoter. Expression of proteins in *E. coli* (T7), baculovirus (p10) and mammalian systems (CMV).	R.S. Eapen
pET19M	Ampicillin	Low-copy number vector for better control by the lac repressor, for N-terminal His_6-tagged recombinant proteins with a TEV protease cleavage site between His_6-tag and recombinant protein. Expression is controlled by the T7 promoter.	Rahman lab, Dept. Pharmacology

pU1	Chloramphenicol	Derivative of pUCDM vector. High-copy number vector, with R6Kγ origin of replication. Designed for expression of multiple open-reading frames for downstream baculovirus expression under control by the p10 or polyhedrin promoters.	Barford lab, LMB [200]
pcDNA3.1(-)	Ampicillin	Plasmid for expression of constructs in mammalian cells under control by the human cytomegalovirus (CMV) immediate early enhancer and CMV promoter.	Invitrogen

Table 2.1: Vectors used for expression in *E. coli*, baculovirus and mammalian cell cultures.

Strain	Resistance	Description	Source
DH5α	-	K-12 strain derivative for routine molecular cloning. High transformation efficiency cells purchased, and chemically competent cells were propagated in-house.	NEB
C41 (DE3)	-	BL21 derived for high-level expression of recombinant proteins. Induction with IPTG induces the expression of T7 RNA polymerase.	Kommander lab, LMB
PIR1	-	Used for molecular cloning and maintenance of plasmids containing an R6Kγ origin.	Barford lab, LMB
DH10 MultiBac^{cre}	Ampicillin, Kanamycin, Tetracyclin, Zeocin	Used to produce Bacmid for insect cell transfection. Stably expressing Cre-recombinase for incorporation of plasmids containing loxP sites into Bacmid genome.	Barford lab, LMB

Table 2.2: *E. coli* strains

2.2 Molecular cloning

2.2.1 Preparation of chemically competent *E. coli*

E. coli cells from a master glycerol stock were streaked onto a LB-Agar plate and incubated overnight at 37°C. The following day, a single colony of cells was picked into 5 mL 2xYT broth and cultured overnight with shaking at 37°C. The following day, 2.5 mL of overnight culture was diluted into 250 mL of LB medium and incubated at 37°C until an OD_{600} of 0.25-0.30 was achieved. The cell culture was split into 50 mL falcons and cooled on ice for 10 minutes. Cell suspensions were pelleted at 4000 × g for 5 minutes and supernatants were carefully removed. Cell pellets were resuspended and pooled in 50 mL TfbI and incubated on ice for 5 minutes prior to repeated centrifugation. The supernatant was removed and the pellet was resuspended in 5 mL of TfbII and incubated on ice for 15 minutes. Competent cells were aliquotted in 50 µL in 2 mL tubes and were stored at -80°C until required.

A modified protocol was used for the preparation of DH10 MultiBac^{cre} cells. Cells were streaked onto a LB-Agar plate with 100 µg/mL ampicillin, 50 µg/mL kanamycin, 10 µg/mL tetracyclin and 25 µg/mL zeocin and grown overnight at 37°C. The following day, a single colony was picked into a 5 mL overnight culture with the above antibiotics. The following day, 2.5 mL of the overnight culture was added into a 250 mL of LB medium, supplemented with 100 µg/mL ampicillin, 50 µg/mL kanamycin, 10 µg/mL tetracyclin and 25 µg/mL zeocin. The culture was grown with shaking at 37°C to OD_{600} of 0.25 before induction with 1 µg/mL L-arabinose to induce the expression of cre-recombinase for 1 hour (final $OD_{600} = 0.50$) prior to harvesting and chemical competent cell preparation as described above.

2.2.2 Polymerase chain reaction (PCR) methods

Two main strategies to obtain the desired DNA construct were used throughout the PhD research. Q5 polymerase (NEB) DNA polymerase was chosen due to its high-fidelity and high speed of amplification. Primers were designed for exponential amplification of the PCR product. All primers were designed with the melting temperatures (T_m) of annealing base pairs to fall between 62-67°C, which were calculated using an online tool provided by NEB [201]. In addition, the T_m of all primer pairs were designed to be within 1-2°C of each other with the 3'-end base pair ending with cytosine (C) or guanine (G).

Standard restriction endonuclease (RE) cloning

This method was primarily used in the transfer of genes from one vector to another. Oligonucleotides were designed with restriction enzyme cleavage sites suitable for the

destination vector at the 5'-end of the primer. PCR reactions were set up in 50 μL reaction volumes and thermocycling was conducted according to the manufacturers' instructions.

Round-The-Horn (RTH) mutagenesis

The second methodology, so called Round-The-Horn (RTH) mutagenesis was primarily used for the addition, deletion or mutation of base pairs at any site within the vector of interest, as depicted in Figure 2.1. Crucially, this methodology does not require the use of restriction enzymes, thus unwanted and additional amino acids encoded by the restriction enzymes are not incorporated into the open reading frame (ORF). All primers were designed end-to-end at the site of modification. Primers used in this method were first 5'- phosphorylated using T4 Polynucleotide Kinase (PNK). Oligonucleotides were dissolved to 100 μM in MilliQ H_2O. Phosphorylation reactions were prepared with 2 μL of 10× PNK Buffer A, 2 μL of 10 mM ATP, 15 μL of 100 μM primer, 1 μL of 10 U/μL PNK and incubated at 37°C for 20 minutes followed by heat inactivation at 75°C for 10 minutes. Phosphorylated primers were then diluted to 10 μM as a working stock solution. PCR reactions were set up in 50 μL reaction volumes and thermocycling was conducted according to the manufacturers' instructions.

2.2.3 Agarose gel electrophoresis

Molecular grade agarose was melted in Tris-Acetate-EDTA (TAE) buffer to an appropriate weight-to-volume percentage dependant on the separation required. SYBR safe DNA stain was added to the molten agarose to a 1× final concentration. Gels were loaded with samples and appropriate base pair ladder(s). Gels were run at 120 V in 1× TAE buffer for 20-35 minutes depending on the separation required before visualising under a UV-transilluminator.

2.2.4 Restriction endonuclease digests and ligations

All restriction endonuclease digests were performed using FastDigest (FD) enzymes and buffer system (Thermo Fisher Scientific) according to manufacturers' instructions. Products from RTH mutagenesis were incubated with FD DpnI to digest parental vector directly in the Q5 polymerase buffer. Vectors prepared for insertion of a PCR product were 5'- and 3'- dephosphorylated using Fast Alkaline Phosphatase (FastAP, Thermo Fisher Scientific) to reduce re-ligation efficiency without the desired insert. Prior to ligation, all DNA were purified using PCR clean up or gel extraction protocols (Qiagen). DNA ligations were performed using Anza™ T4 DNA ligase Master Mix (Thermo Fisher Scientific) with a downscaled protocol from 20 μL to 4 μL reaction volumes at room temperature for 30 minutes. Ligations with an insert and vector were prepared with a 3:1 molar ratio,

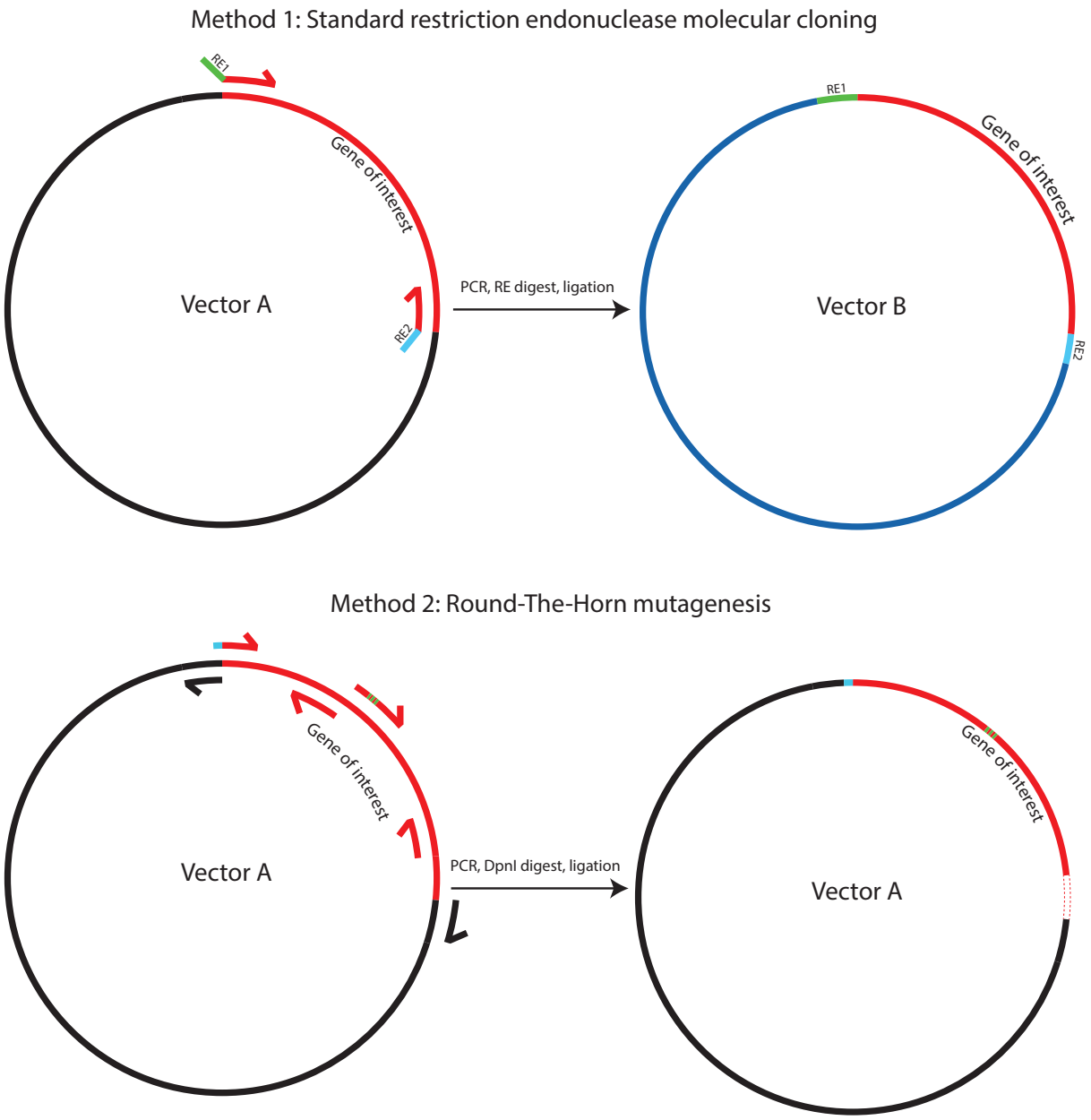


Figure 2.1: Schematic of molecular cloning methods. Method 1: Standard restriction endonuclease (RE) cloning used to transfer genes between two vectors. RE sites coloured in green and yellow. Method 2: Examples of three primer pairs, used for the addition (cyan), mutagenesis (green) and deletion (red-dashed) of base pairs.

not exceeding 100 ng of vector. Ligations from RTH PCR did not exceed 100 ng of PCR product.

2.2.5 Transformation into chemically competent cells

Plasmid DNA (10-50 ng) were used for the transformation of whole plasmids and for the transformation of ligation reactions, 4 µL (not exceeding 100 ng vector) were used. The relevant *E. coli* strain for the downstream application was thawed on ice. Plasmid or

ligation reaction was added to the cells, giving a gentle flick to mix the contents and was then incubated on ice for 30 minutes. *E. coli* were heat shocked at 42°C for 35-45 seconds before returning immediately to ice for 2 minutes. SOC medium (200 µL, NEB) was added and cells were recovered at 37°C, shaking at 800 rpm for 30 minutes (ampicillin) or 60 minutes (kanamycin and chloramphenicol). The cell suspension was then spread onto LB-Agar plates with the appropriate antibiotics, using sterile glass beads. LB-Agar plates were then incubated overnight at 37°C.

Transformations into DH10 MultiBac^{cre} were performed with a modified protocol. A 50 µL aliquot of DH10 MultiBac^{cre} was thawed on ice and 1 µg of pU1 vector was added. The cells were incubated on ice for 30 minutes and heat shocked at 42°C for 40 seconds. The cells were then placed on ice for 2 minutes before diluting into 6× 0.5 mL of 2xYT medium. Cells were recovered overnight at 30°C, shaking at 800 rpm. The following day, cells were pooled and centrifuged at 500 × g for 1 minute prior to resuspending in 250 µL 2xYT medium. The cell suspension was spread on a LB-Agar plate supplemented with 100 µg/mL ampicillin, 50 µg/mL kanamycin, 34 µg/mL chloramphenicol, 40 µg/mL IPTG and 100 µg/mL X-gal. The LB-Agar plate was incubated at 37°C for approximately 72 hours until positive (blue) colonies were observed.

2.2.6 DNA purification from recombinant source

Plasmid preparation

Single colonies from transformation plates were picked into 5 mL 2xYT medium supplemented with the appropriate antibiotics. Cultures were grown overnight at 37°C, shaking at 180 rpm. The following day, 2 mL of overnight culture was pelleted in a 2 mL tube. The supernatant was decanted and the cell pellet was resuspended by vortex in 250 µL of P1 resuspension buffer (Qiagen) with 100 µg/mL RNase A (Qiagen). P2 buffer (250 µL) was added and tubes were inverted 5 times to mix, followed by the addition of 350 µL of N3 buffer. The resulting suspension was inverted several times to mix thoroughly prior to centrifugation at 20,000 × g for 15 minutes. The resulting supernatant (700 µL) was applied to a DNA spin column (Expedeon) and spun through, discarding the flow-through. The spin column was washed with 700 µL DNA wash buffer, discarding the flow-through. Centrifugation was repeated to remove residual wash buffer. Bound DNA was eluted using 50 µL of MQ H_2O into a 1.5 mL tube. DNA concentrations were measured using a Nanodrop 2000 (Thermo Fisher Scientific) against a blank of MQ H_2O. Plasmid DNA was stored at -20°C until required. All plasmids were Sanger sequenced by Eurofins Genomics before use in downstream experiments.

Bacmid preparation

A single blue colony from a transformed DH10 MultiBac^{cre} LB-Agar plate was picked into 10 mL of 2xYT medium supplemented with 100 μg/mL ampicillin, 50 μg/mL kanamycin, 34 μg/mL chloramphenicol and grown overnight at 37°C, shaking at 180 rpm. The following day, cells were pelleted at 7000 × g for 2 minutes, discarding the supernatant. The cell pellet was resuspended in 500 μL of P1 buffer in a 2 mL tube. P2 buffer (500 μL) was then added and tube was inverted several times, gently. N3 buffer (700 μL) was added and the tube was inverted until the suspension was thoroughly mixed before centrifugation at 20,000 × g for 10 minutes. Bacmid DNA was precipitated by the addition of 1 mL of ice-cold isopropanol to 1 mL of the resulting supernatant. Precipitated DNA was then pelleted at 20,000 × g, 4°C for 5 minutes and the supernatant was carefully removed. The DNA pellet was washed with 1 mL of 70% (v/v) ethanol and was centrifuged once more before removing the supernatant. The DNA pellet was allowed to air-dry briefly before dissolving in 75 μL of MQ H_2O. DNA concentrations were measured using a Nanodrop 2000 (Thermo Fisher Scientific) against a blank of MQ H_2O. Bacmid DNA was stored at -20°C until required. All bacmids were sequenced by Eurofins Genomics before use in downstream experiments.

2.3 Protein expression and purifications

Methods describing protein purifications for Tankyrase and Cdc20 domains are included in chapters 3 and 4, respectively. This section describes the expression and purification of proteins used downstream in the above projects and for collaborative work conducted on Mdm2, Appendix C. Amino acid sequences for all proteins are provided in Appendix A.

2.3.1 Analysis of proteins

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS PAGE)

Gels were cast using Biorad mini-casting cassettes. Cassettes were constructed with appropriate spacers depending on sample load volume. Separating gel solutions were prepared with separating buffer, MQ H_2O, ammonium persulphate (APS), tetramethylethylenediamine (TEMED) and the required percentage of acrylamide/ bis-acrylamide solution (37.5:1 ratio). The gel was cast into the cassette and levelled with a layer of isopropanol (IPA). Once set, the IPA was removed and residual solvent was dried off. Stacking gel was prepared with stacking buffer, MQ H_2O, APS, TEMED and 5.1 % (w/v) acrylamide/bis-acrylamide. Stacking gel were added and a 15 or 10-well comb was used to set the wells. Spectra^{TM} Multicolour Broad Range protein ladder (Thermo Fisher Scientific) was loaded

on all gels. Gels were run in Tris-Glycine-SDS buffer at 120 V (constant) for 10 minutes followed by 50-65 minutes at 180 V (constant), depending on the percentage of acrylamide. Gels were removed from the glass plates and were washed thrice in hot tap water to remove SDS from the gel. Gels were then stained in Coomassie brilliant blue for at least 30 minutes and destained in water. Gels were subsequently imaged using a Gel-doc (UVP) under white light.

Electrospray-ionisation mass spectrometry (ESI-MS) of proteins

Selected proteins were mass verified by ESI-MS using a Xevo G2-S QTOF (Waters) instrument, in the mass spectrometry facility, Department of Chemistry, University of Cambridge. All samples were injected onto an Acquity Protein BEH 300A, 1.7 μm, 2.1×50 mm, UPLC column (Waters) connected to an Acquity UPLC (class I) system (Waters). The liquid chromatography used HPLC grade H_2O, 0.1% (v/v) formic acid (A) and MeCN, 0.1% (v/v) formic acid (B) at 0.2 ml/min in a linear gradient from 95:5 (A:B) to 0:100 (A:B) over four minutes, before returning to original conditions. The Lockspray™ was set with Leucine enkephalin and was injected at regular intervals during the run as a known reference mass. Data collection and processing were performed using MassLynx 4.1. Protein masses were calculated using the maximum entropy function with appropriate boundaries set for the expected molecular weight of the protein. Calculations were allowed to iterate until data no longer improved.

Calculation of recombinant protein concentration

Full amino acid sequences for all recombinant proteins produced were known and extinction coefficients at 280 nm (ϵ_{280}) were estimated using ExPASy's ProtParam server [202]. Absorbance at 280 nm was then measured against a buffer blank using a Nanodrop 2000 (Thermo Fisher Scientific) and molar concentrations calculated according to Beer-Lambert's law, equation 2.1

$$A = \epsilon c l \tag{2.1}$$

where, A is the absorbance at a given wavelength, ϵ is the molar extinction coefficient at the same wavelength, c is the molar concentration and l is the path length in cm.

For the purpose of crystallography, where protein is more commonly dealt with in units of mg/mL, molar concentrations were converted according to the protein's estimated molecular weight.

Calculation of total protein concentration

Total protein concentrations of mammallian cell lysates were calculated using Pierce Bicinchoninic acid (BCA) assay kit (Thermo Fisher Scientific). BCA assays were performed as per manufacturers' instructions. Experimental sample concentrations were determined against a standard curve generated from serial dilutions of 2 mg/mL bovine serum albumin, supplied with the kit.

2.3.2 Tandem Immobilised Metal ion Affinity Chromatography (IMAC)

Affinity chromatography is a commonly used technique to obtaining relatively pure protein from the crude source. The His_6-tag is a convenient tag due to its small size and its affinity to Ni^{2+}. Ni^{2+} ions can be immobilised to solid support matrices such as agarose beads or similar resins to aid chromatographic separation of His_6-tagged biomolecules. Some proteins do not tolerate the elution conditions, which can be lower pH ($<$ pH 7.0) resulting in protonation of the His_6-tag or more commonly the addition of high concentrations of imidazole to outcompete the imidazole rings of the His_6-tag. To minimise the denaturation of proteins during IMAC elution due to the presence of imidazole, a new method was developed. This method relied on the direct and immediate desalting of the IMAC eluent thereby minimising the contact time of proteins with imidazole-containing buffers. This was achieved by coupling the outlet from the IMAC column directly to the inlet of a pre-equilibrated 26/10 desalting column (Figure 2.2). A maximum of 25 mL IMAC elution can be applied to the desalting column to achieve optimal separation. Proteins then elute from both columns, with the protein immediately desalted into the buffer of choice.

2.3.3 Tobacco Etch Virus (TEV) protease (S219V)

The gene for N-terminally His_6-tagged TEV protease was provided in an unknown vector (Itzhaki lab). The S219V mutation was introduced into the gene by RTH-mutagenesis resulting in a construct that is less prone to self-cleavage [203]. The plasmid encoding for His_6-TEV protease (S219V) was transformed into C41 cells and grown in 2xYT medium supplemented with 100 μg/mL ampicillin at 37°C, shaking at 200 rpm to an OD_{600} of 0.6 - 0.7 in 2 L baffled erlenmeyer flasks. Cells were induced with 200 μM IPTG overnight at 20°C, shaking at 200 rpm. The following day, cells were pelleted at 7,000 × g for 15 minutes. For every 2 L of culture, cell pellets were resuspended in 40 mL of lysis buffer (50 mM Tris-HCl, 300 mM NaCl, 20 mM imidazole, 1 mM TCEP, pH 8.0) and stored at -20°C until purification was required. Resuspended cells were thawed at room temperature, and crude DNase I (Sigma Aldrich) was added to reduce the viscosity during

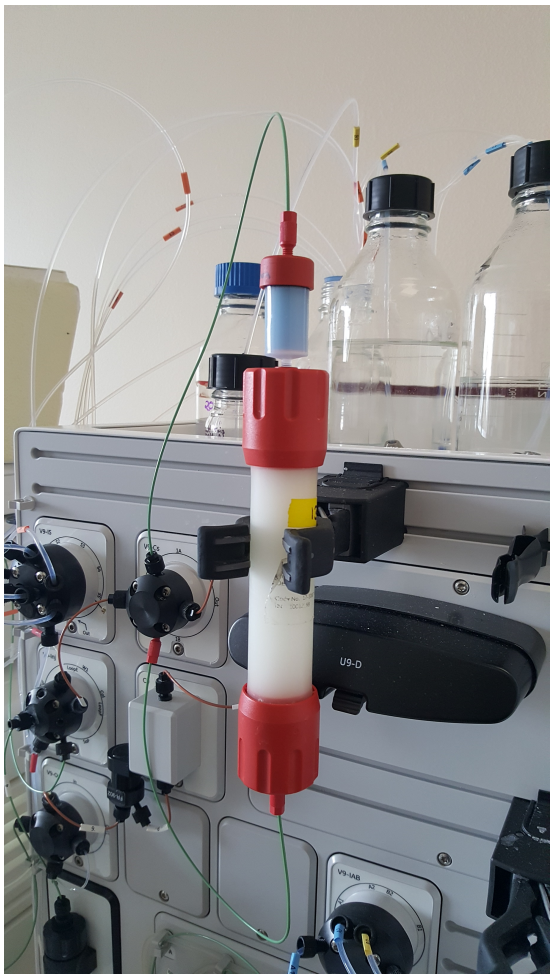


Figure 2.2: Chromatography setup on ÄKTA Pure system for tandem IMAC-desalt of proteins elutions.

lysis. The cell suspension was lysed by 3 passes on a C5-Emulsiflex (Avestin) at 15,000-20,000 psi. The lysate was clarified by centrifugation at 45,000 × g, 4°C for 45 minutes. All chromatography steps were performed using an ÄKTA Pure with a sample pump. First, a 26/10 desalting column was equilibrated into desalting buffer (50 mM Tris-HCl, 300 mM NaCl, 1 mM TCEP, pH 8.0) and disconnected from the system. Next, a 5 mL HisTrap Excel column was pre-equilibrated in wash buffer (50 mM Tris-HCl, 300 mM NaCl, 20 mM imidazole, 1 mM TCEP, pH 8.0) prior to loading the clarified supernatant at 3 mL/min. The column was washed with 20 column volumes (CV) of wash buffer to remove contaminant proteins. Proteins were then eluted as described in 2.3.2 using elution buffer (50 mM Tris-HCl, 300 mM NaCl, 300 mM imidazole, 1 mM TCEP, pH 8.0) into a 96 deep-well block in 2 mL fractions. Elutions were pooled and concentrated using a Vivaspin 20, 10K Molecular weight cut off (MWCO) centrifugal filter at 8000 × g, 4°C until a concentration of 270 µM was achieved (ϵ_{280} = 33460 $M^{-1}cm^{-1}$). Protein was then diluted with an equal volume of 100 % glycerol, resulting in 135 µM protein in 25 mM Tris-HCl, 150 mM NaCl, 0.5 mM TCEP, 50 % (v/v) glycerol, pH 8.0. Proteins were

analysed by SDS PAGE and estimated to be >75% pure (Figure 2.3). Due to its use in early stage purifications and at relatively low concentrations, this purity was considered sufficient. Protein was then aliquotted, flash frozen in liquid N_2 and stored at -80°C until required.

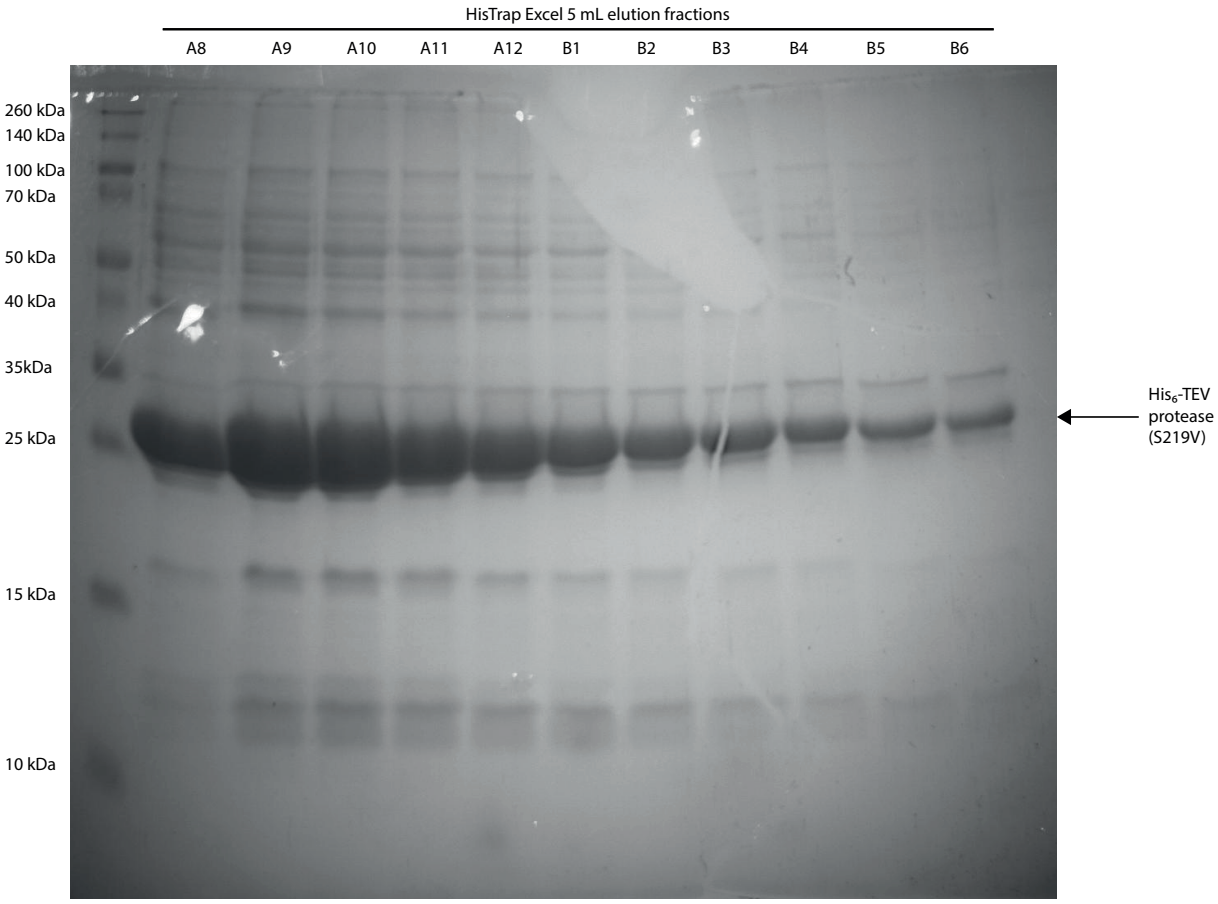
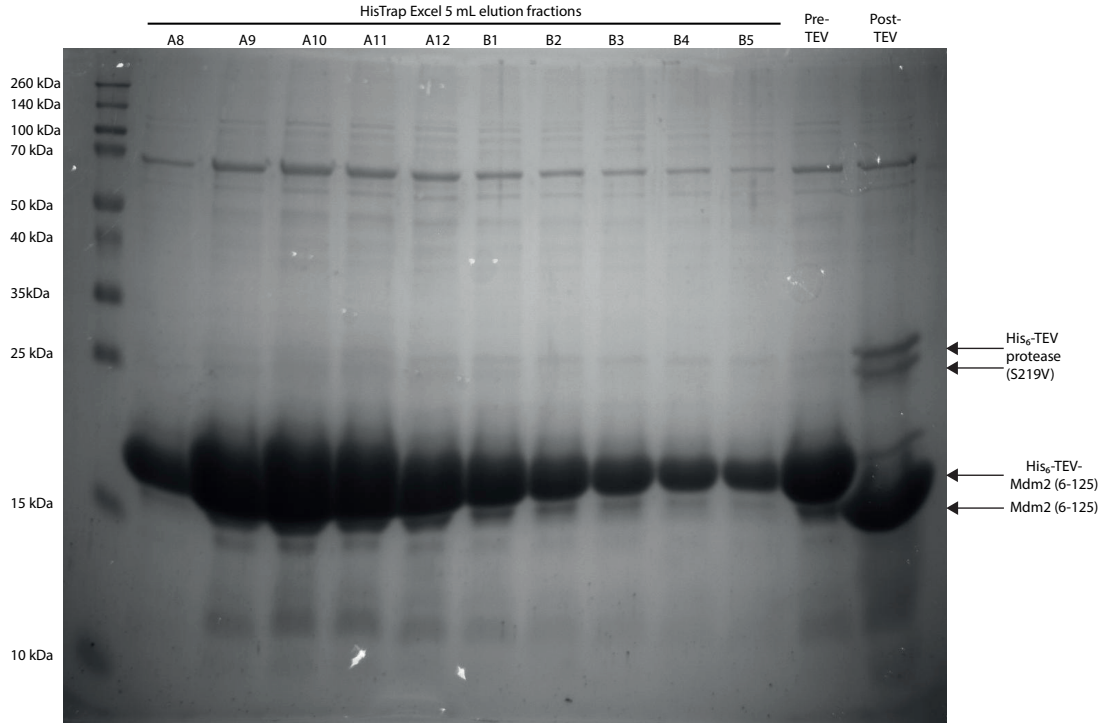


Figure 2.3: SDS PAGE gel of His_6-TEV-protease (S219V) elution fraction from IMAC purification (lanes A8-B6). His_6-TEV-protease (S219V) appears to migrate slightly faster through the SDS PAGE gel than its expected molecular weight of 28.6 kDa.

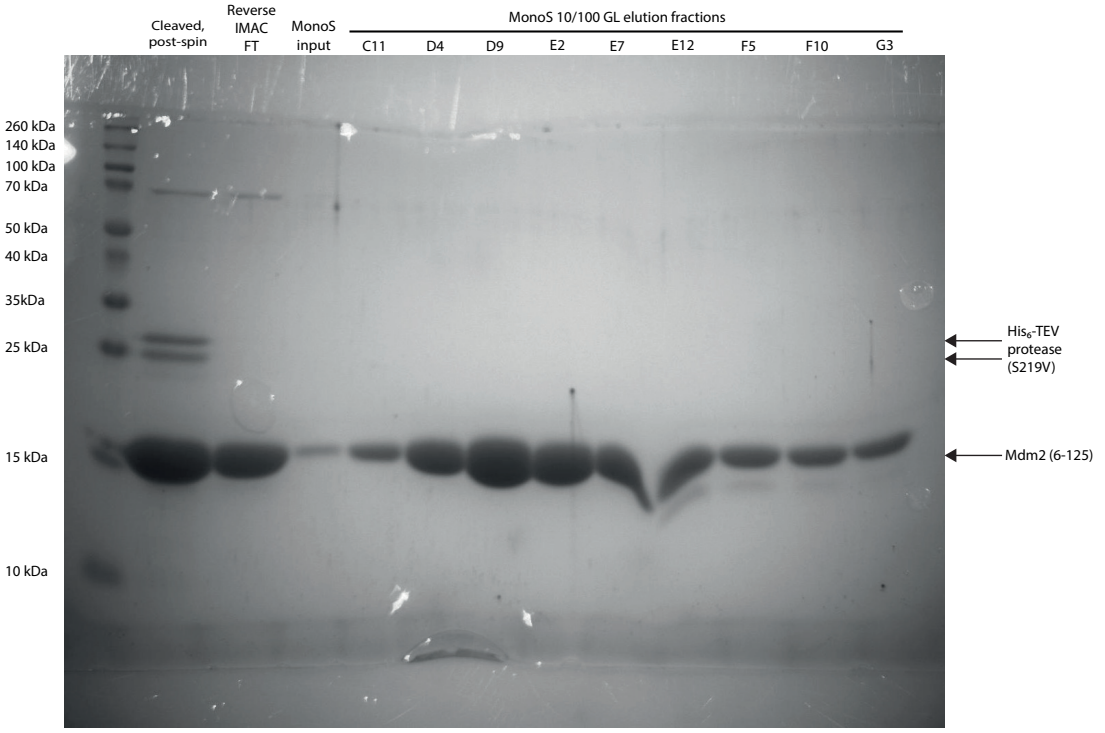
2.3.4 Mouse double minute 2 (Mdm2)

The gene for human Mdm2, residues 6-125 was cloned into the pRSET-TEV vector. The ORF encodes a N-terminal His_6-tag, followed by a 10 residue asparagine linker and TEV-cleavage site followed by Mdm2 (6-125). The plasmid was transformed into C41 cells and grown in 2xYT medium supplemented with 100 μg/mL ampicillin at 37°C, shaking at 200 rpm to an OD_{600} of 0.6 - 0.7 in 2 L baffled erlenmeyer flasks. Cells were induced with 300 μM IPTG overnight at 18°C, shaking at 200 rpm. The following day, cells were harvested by centrifugation at 7000 × g for 15 minutes. The cell pellet from 2 L of cell culture was resuspended in 25 mL ice-cold lysis buffer (50 mM Tris-HCl, 300 mM

NaCl, 15 mM imidazole, 1 mM DTT, SigmaFAST EDTA-free protease inhibitor cocktail tablet (Sigma-Aldrich) and crude DNase I (Sigma-Alrdich), pH 8.0). The cell suspension was lysed by 3 passes on a C5-Emulsiflex (Avestin) at 15,000-20,000 psi. The lysate was clarified by centrifugation at 45,000 × g, 4°C for 45 minutes. All chromatography steps were performed using an ÄKTA Pure with a sample pump. First, a 26/10 desalting column was equilibrated into desalting buffer (20 mM Tris-HCl, 180 mM NaCl, 1 mM DTT, pH 8.0) and disconnected from the system. Next, a 5 mL HisTrap Excel column was pre-equilibrated in wash buffer (50 mM Tris-HCl, 300 mM NaCl, 15 mM imidazole, 1 mM DTT, pH 8.0) prior to loading the supernatant at 3 mL/min. The column was washed with 20 column volumes (CV) of wash buffer to remove contaminant proteins. Proteins were then eluted as described in 2.3.2 using elution buffer (50 mM Tris-HCl, 300 mM NaCl, 300 mM imidazole, 1 mM DTT, pH 8.0) into a 96 deep-well block in 2 mL fractions. Elutions were pooled and His_6-TEV protease (S219V, 0.5 μM final concentration) was added to remove the His_6-tag from Mdm2 overnight at 4°C. The following day, precipitates were removed by brief centrifugation and supernatant was flowed over a 5 mL HisTrap Excel column at 3 mL/min, collecting the flow-through. The flow through was diluted to a final concentration of 30 mM NaCl in MonoS low salt buffer (LSB; 20 mM Bis-Tris, 1 mM DTT, pH 6.5). A MonoS 10/100 GL column was equilibrated in 20 mM Bis-Tris, 30 mM NaCl, 1 mM DTT, pH 6.5 prior to loading the diluted protein sample from above at 4 mL/min. Bound proteins were then eluted at 2 mL/min over 20 CV in a linear gradient with MonoS high salt buffer (HSB; 20 mM Bis-Tris, 1 M NaCl, 1 mM DTT, pH 6.5). Mdm2 eluted in a sharp peak around 200 mM NaCl and fractions were analysed by SDS PAGE prior to being pooled and concentrated in a Vivaspin 20, 5K MWCO centrifugal filter at 8000 × g, 4°C. Protein concentration was determined to be 195 μM (ϵ_{280} = 10,430 $M^{-1}cm^{-1}$) and was estimated to be >95% pure by SDS PAGE analysis (Figure 2.4). Mdm2 (6-125) was also verified to be the correct expected mass by ESI-MS. Proteins were flash frozen in liquid N_2 and stored at -80°C.



(a) Mdm2 IMAC purification



(b) Mdm2 IEX purification

Figure 2.4: SDS PAGE gels monitoring the purification process. (a) IMAC elutions of proteins bound to the HisTrap Excel resin. TEV-cleavage efficiency is monitored taking fractions before and after cleavage. (b) Most contaminant proteins precipitate out of solution and could be separated by centrifugation. MonoS purification cleans up remaining major contaminants.

2.4 Software

GraphPad Prism

Graphs were plotted and data were fitted using GraphPad Prism version 7.00 for Windows, GraphPad Software, La Jolla California USA, www.graphpad.com. Data were fitted using in-built mathematical models, where available. For all other fits, data were fitted using equations described in the relevant method section of the chapter. Unless otherwise specified, error bars are plotted as the standard error (SE).

ChemDraw Ultra 12.0

Structures of synthesised peptides were drawn using ChemDraw Ultra 12.0 (CambridgeSoft, PerkinElmer).

PyMOL

Open-source PyMOL (Schrödinger, LLC, latest version updated as it is released) [204] was installed using Python 3.6 using a pre-compiled .whl file for Windows operating systems by Christoph Gohlke [205]. PyMOL additionally requires the NumPy+MKL library and Pmw to function, which were also obtained from Golhke's website. All models and images of protein structures were produced using PyMOL software.

CCP4i

CCP4i (version 7.0.077) was installed for Windows OS [206]. Subsiduary programs contained within CCP4i, which were also used include Matthews coefficient [207, 208], Phaser [209], Refmac5 ([210] and references therein) and PDBset.

BUSTER

BUSTER (Global Phasing) version 2.10.3 was used for X-ray structure refinement [211, 212].

WinCoot

Coot (version 0.8.9.2) [213] for Windows OS was installed from Bernhard Lohkamp's pre-compiled installer [214].

Chapter 3

Targeting the Tankyrase substrate-binding domains using constrained peptides

3.1 Introduction and aims

The work presented in this chapter was conducted in the first year of the PhD. The project was aimed at developing improved peptide-based inhibitors against the substrate recruitment domains of TNKS. The project largely built on the work carried out by a former PhD student (Dr. Wenshu Xu), who had screened several iterations of macrocyclisation with the double-click stapling technologies developed in the Spring lab [99]. While the lead peptides developed were cell permeable and able to inhibit WNT signalling, their affinities were only 2-fold improved in comparison to the optimised sequence by Guettler *et al.* [131]. However, improvements of the peptide binding affinity upon constraining were not expected due to the extended nature of the binding interface. The initial goal was to improve the peptide binding affinity to TNKS, which in turn should subsequently lower the IC_{50}, yielding a more potent inhibitor.

3.2 Methods

3.2.1 Expression and purification of wild-type & mutant Tankyrase ARC domains

His_6-tagged TNKS1 ARC2-3 (residues 315-662) and His_6-tagged TNKS2 ARC4 (residues 488-649) were expressed and purified using the same protocols. Wild-type and mutant GST-tagged TNKS2 ARC4 were expressed and purified using the same protocols. C41 cells containing a transformed TNKS gene were grown to $OD_{600} \approx 0.6$, in a shaking

incubator at 37°C prior to induction with 150 μM IPTG at 20°C overnight. Cells were harvested by centrifugation at 4000 × g, 4°C. Pellets were resuspended in 20 mL of lysis buffer (50 mM Tris-HCl, 500 mM NaCl, 2 mM DTT, 0.1% (v/v) Tween 20, and one SigmaFAST EDTA-free Protease inhibitor cocktail tablet, pH 8.0) per 0.5 L of culture. Pellets of His_6-tagged proteins also included 10 mM imidazole. Cells were incubated with lysozyme and DNase I at 4°C on a roller for at least 30 minutes prior to lysis by sonication on ice. Lysed cells were clarified by centrifugation at 40,000 × g, 4°C for 40 minutes. Supernatants were bound to 5 mL of Ni-NTA or glutathione-agarose resin (Expedeon) in a gravity-flow column, for 30 minutes at 4°C. Resins were washed with 250 mL of wash buffer (50 mM Tris-HCl, 500 mM NaCl, 2 mM DTT, pH 8.0 (+ 10 mM imidazole)). Bound proteins were eluted with elution buffer (50 mM Tris-HCl pH 8.0, 500 mM NaCl, 2 mM DTT) containing either 500 mM imidazole (His_6-tagged protein) or 20 mM reduced glutathione (GST-tagged protein). Proteins were subsequently pooled and desalted into 50 mM Tris-HCl pH 8.0, 300 mM NaCl, 2 mM DTT using Centripure P25 columns (Generon) according to the manufacturer. To remove the His_6-tag, 125 U of thrombin (Sigma-Aldrich) were added, incubating overnight at room temperature on a rotating mixer. Cleaved proteins were clarified by brief centrifugation and the supernatant containing His_6-tagged protein were subjected to reverse-IMAC, collecting the flow through. Proteins were concentrated using a 10K MWCO centrifugal filter (Sartorius Stedim) to ≈10 mL. Samples were then separated by SEC using a 26/60 Superdex 75 column in SEC buffer (50 mM HEPES pH 7.5, 300 mM NaCl, 2 mM DTT). Elutions were verified for purity via SDS PAGE and appropriate fractions were pooled and concentrated to $>$ 500 μM as determined by absorbance at 280 nm. For each construct, molecular weight and molar extinction coefficients (cysteines reduced) were determined using ProtParam from the ExPASy web server [202] and are listed in Table 3.1. Proteins were flash frozen in liquid N_2 and stored at -80°C until required.

3.2.2 Peptide synthesis and macrocyclisation of T1-1

Peptide T1 was synthesised on Rink amide resin (Novabiochem) using standard Fmoc-chemistry on a Liberty Blue peptide synthesiser (CEM). Fmoc-L-Ornithine(N_3) was synthesised as described by Lau and Spring [215]. The amino terminus of peptide T1 was then deprotected with 20% (v/v) piperidine in DMF and acetylated in a cocktail of 40% (v/v) acetic anhydride, 40% (v/v) DMF and 20% (v/v) DIEA to cap the peptide. The resin was then washed 3 times in DMF and dried twice in MeOH and then twice in Diethyl Ether. T1 was cleaved from the Rink amide resin in a cocktail containing 95% (v/v) TFA, 2.5% (v/v) H_2O and 2.5% (v/v) TIPS. Eluate was dried under N_2 gas and triturated with diethyl ether. T1 was then purified by reverse-phase HPLC and the mass was verified by

Table 3.1: List of all constructs used in this chapter. Theoretical molecular weight and extinction coefficient at 280 nm are tabulated. All values are calculated using Expasy ProtParam web tool [202].

Protein	Molecular weight (Da)	ϵ_{280} (M^{-1} cm^{-1})
His_6-TNKS1 ARC2-3	39790.47	15930
TNKS1 ARC2-3	38013.47	15930
His_6-TNKS2 ARC4	19717.17	12950
TNKS2 ARC4	17708.97	12950
GST-TNKS2 ARC4	43988.59	55810
GST-TNKS2 ARC4 (G558S)	44018.61	55810
GST-TNKS2 ARC4 (G558T)	44032.64	55810
GST-TNKS2 ARC4 (G558D)	44046.62	55810

LCMS.

Double-click stapling was performed in degassed 50% (v/v) t-BuOH, ddH_2O. Peptide T1 was dissolved at 1 mg/mL. A molar ratio of 1.1:1 of hepta-1,6-diyne linker to T1 was added to the mix. To catalyse the reaction, a 4× stock of sodium ascorbate, $CuSO_4.5H_2O$ and THPTA was dissolved in a 3:1:1 molar ratio in the click chemistry solvent system. The catalyst stock solution was then added to the peptide-linker solution at a 1× final concentration. The reaction was stirred at room temperature under nitrogen for 20 hours. The reaction progress was monitored by LCMS. The final T1-1 peptide was purified by reverse-phase HPLC. The concentration of working stock peptide solution was verified by amino acid analysis (Peter Sharrat, former PNAC facility, Department of Biochemistry, University of Cambridge).

3.2.3 Differential scanning fluorimetry

Experiments were performed on Roche Lightcycler 480 II or BioRad CFX Connect machines in 384-well or 96-well formats, respectively. SYPRO Orange probe was used as the reporter dye molecule at a 5× final concentration. TNKS1 ARC2-3 and TNKS2 ARC4 were prepared at a final assay concentration of 5 μM diluted in SEC buffer, in the presence or absence of ligands, maintaining 1% (v/v) DMSO. Thermal melt profiles were obtained by ramping the block from 20°C to 95°C at a rate of 0.5°C/min. The SYPRO Orange dye was excited at 483 ± 35 nm and the fluorescence was monitored at 568 ± 20 nm. Data were fitted in GraphPad Prism 7.0 using

$$Y = (Y_{min} + M2 \times X) + \frac{(Y_{max} + M1 \times X) - (Y_{min} + M2 \times X)}{1 + e^{\frac{V50 - X}{m}}}, \tag{3.1}$$

where Y is the fluorescence intensity at a given temperature, X, Y_{min} and Y_{max} describe the minimum and maximum fluorescence intensity, respectively, $M2$ and $M1$ are the pre- and post- transition baseline gradients, $V50$ describes the midpoint of the melting transition and m is the gradient of the melting transition.

3.2.4 Fluorescence polarisation

Direct FP experiments

The dissociation constant, K_d, of the fluorescent peptide ligand TAMRA-pep1 was determined against the Tankyrase substrate-binding domains; TNKS1 ARC2-3, TNKS2 ARC4 and GST-TNKS2 ARC4 (G558X) mutants, hereby referred to as TNKS. TNKS were titrated in a 1.5-fold serial dilution against 20 nM TAMRA-pep1 in black 384-well plate (OptiPlate-384 F, Perkin Elmer). Protein and peptide solutions were diluted in PBS+0.01% Tween 20, pH 7.4. Endpoint equilibrium measurements were read after a thirty minute incubation at room temperature using a ClarioStar microplate reader (BMG, Labtech). TAMRA was excited at 540 ± 20 nm and the emmission was measured at 590 ± 20 nm, with a dichroic filter at 566 nm. Polarisation is then calculated using

$$mP = \frac{1000 \times (F_{\parallel} - F_{\perp})}{F_{\parallel} + F_{\perp}}, \tag{3.2}$$

where $F_{\parallel}$ is the fluorescence from vertically polarised light and $F_{\perp}$ is the fluorescence from horizontally polarised light and 1 polarisation unit $(P) = 1000\ mP$.

Dissociation constants of TAMRA-pep1 to TNKS proteins were calculated by fitting the data to

$$r = r_0 + (r_b - r_0) \times \frac{(K_d + [L]_t + [P]_t) - \sqrt{(K_d + [L]_t + [P]_t)^2 - 4[L]_t[P]_t}}{2[L]_t} \tag{3.3}$$

where r is the calculated polarisation, r_0 is the polarisation of free peptide, r_b is the polarisation of the protein-ligand complex, K_d is the dissociation constant of the protein-ligand complex, $[P]_t$ is the total protein concentration and $[L]_t$ is the total ligand concentration.

Competition FP experiments

In competition fluorescence polarisation experiments, unlabelled peptides competing for the same binding site were titrated against a TNKS-TAMRA-pep1 complex. Since the dissociation constant, K_d of the TNKS-TAMRA-pep1 complex is known, it is possible to determine the dissociation constants for unlabelled ligands using [216]

$$r = r_0 + (r_b - r_0) \times \frac{2\sqrt{d^2 - 3e}\cos\frac{\theta}{3} - d}{3K_{d1} + 2\sqrt{d^2 - 3e}\cos\frac{\theta}{3} - d}, \tag{3.4}$$

where d is described as

$$d = K_{d1} + K_{d2} + [L]_{st} + [L]_t - [P]_t, \tag{3.5}$$

and e is described as

$$e = ([L]_t - [P]_t)K_{d1} + ([L]_{st} - [P]_t)K_{d2} + K_{d1}K_{d2}, \tag{3.6}$$

and θ is

$$\theta = \arccos \frac{-2d^3 + 9de - 27(-K_{d1}K_{d2}[P]_t)}{2\sqrt{(d^2 - 3e)^3}}. \tag{3.7}$$

Parameters in Equation 3.4 are the same as in Equation 3.3, with the exception of $[L]_t$, which now describes the total concentration of competing peptide and $[L]_{st}$, which describes the total concentration of TAMRA-pep1. The dissociation constant of the TNKS-TAMRA-pep1 complex, K_d, is now referred to as K_{d1}, and the dissociation constants of unlabelled competing peptides to TNKS is now referred to as K_{d2}.

3.2.5 Surface plasmon resonance

Experiments were run using a Biacore T200 instrument (GE healthcare) using a Series S, Ni-NTA chip (GE healthcare). Experiments were conducted in the following running buffer; 50 mM HEPES, 150 mM NaCl, 2 mM TCEP, 0.005% (v/v) Tween 20, pH 7.5. To maintain continuity of nomenclature in the SPR field, proteins are hereby considered the ligand or receptor, whereas peptides are considered analytes.

Ligand immobilisation

The Ni-NTA chip was docked and primed three times in running buffer. Each flow cell was processed in succession. First, flow cells were stripped of Ni^{2+} with 350 mM EDTA, followed by re-equilibration with running buffer. The flow cell was then recharged with 500 µM $NiSO_4$. EDC and NHS were then mixed and immediately flowed over to activate the surface for covalent coupling. 2 µM His_6-TNKS2 ARC4 (Fc 2), 10 µM His_6-TNKS2 ARC4 (Fc 3) and 200 nM His_6-TNKS1 ARC2-3 (Fc 4) were flowed over their respective flow cells with a ligand target capture set at 4000 RU (Fc 2, 3) and 6000 RU (Fc 4). All flow cells were blocked using 100 mM Tris-HCl, pH 8.0. Non-covalently bound ligands were removed with an injection of 350 mM EDTA.

Analyte binding experiments

Peptides that are previously known to bind TNKS ligands, were run as positive control molecules. Stock solutions at 10 µM of pep1, TAMRA-pep1 and sp4n2m3-AntP were titrated from in a 2-fold serial dilution against the immobilised ligands with 120 second association and 180 second dissociation times.

3.2.6 Isothermal titration calorimetry

All experiments were performed using a MicroCal ITC200 (Malvern panalytical) at 25°C. Proteins were dialysed in 50 mM Tris-HCl, 150 mM NaCl, 2 mM TCEP, pH 7.5. Peptide was diluted in the dialysis buffer and solutions were appropriately 'buffer matched'. Twenty injections of 2 µl were performed over 4 seconds each, spaced by 120 seconds between injections. Data collection was filtered to 2 seconds. Reference power was set at 6 µcal/sec, with a a spin speed of 800 rpm. Data were then fitted in Origin 7.0 software with a one-site binding model.

3.2.7 X-ray crystallography

Commercial or in-house screens developed by AstraZeneca were dispensed into 96-well MRC 2-drop plates. For custom grid screens, plates were dispensed using a Dragonfly crystal (TTP Labtech). Dispensed solutions were then mixed for 5 minutes using a MXone (TTP Labtech). His_6-TNKS2 ARC4 or TNKS2 ARC4 were used in crystal trials at 20 mg/ml. Proteins and well solutions were dispensed using a Mosquito HTS (TTP Labtech). Plates were sealed with Crystal clear sealing tape (Hampton Research) and stored at 20°C in a Rock Imager (Formulatrix). Visible light and UV images were collected along the incubation period to monitor each drop.

3.3 Results

3.3.1 Cloning and expression of wild-type Tankyrase proteins

Plasmids encoding for TNKS1 ARC2-3 and TNKS2 ARC4 were provided by Dr. Wenshu Xu. TNKS1 ARC2-3 was encoded in a modified pRSETa vector with an N-terminal His_6 tag, whereas TNKS2 ARC4 was provided in a modified pGST vector. To ease the processivity and purification of both recombinant proteins side-by-side, the TNKS2 ARC4 ORF was transferred by standard restriction-endonuclease cloning into a modified pOPIN F vector (Dr. Owen Burbidge) containing an N-terminal His_6 tag. For the work exploring the binding cavity around Arg1 of the TNKS-binding peptides, mutations were incorporated in GST-TNKS2 ARC4 at G558 to S, T and D by Round-the-Horn SDM.

Expression of all constructs were very high, typically yielding $>$ 50 mg of purified protein per litre of induced *E.coli* culture. Proteins were either purified by IMAC or GST-glutathione affinity chromatography. For wild-type proteins affinity tags were typically removed by thrombin cleavage for downstream biophysical or crystallographic applications and further purified by Reverse IMAC (His_6-tagged proteins only) and size-exclusion chromatography. The GST-TNKS2 ARC4 (G558X) proteins were characterised without

removal of the GST-tag. All proteins were deemed >95% pure, analysed by SDS PAGE.

3.3.2 Biophysical method development

We wanted to expand the biophysical assays that Dr. Wenshu Xu had previously set up to assess peptide binding to the TNKS proteins (FP and ITC). We initially deemed SPR to be a highly useful tool for this purpose due to the ability to extract kinetic information from the system. This data could be very useful in building up a SAR with the goal of improving peptide binding affinity. Although we were able to immobilise sufficiently high RU of our protein ligands to observe analyte signals at ≈100 RU, we observed particularly poor activity of the ligand after immobilisation. Although both pep1 and TAMRA-pep1 were able to achieve a similar response, it was very clear that sp4n2m3-CPP was binding to the reference flow cell significantly. This is likely to be due to an interaction between the net-positively charged Antennapedia cell-penetrating peptide to the negatively charged dextran surface of the Ni-NTA sensor chip. We sought to investigate the reason behind the lack of activity on the chip despite the high density of ligand on the chip surface. We used ITC as our first method as it would give an indication to the ability of the reagents to bind each other in a stoichiometric manner.

The data presented in Figure 3.1 highlights that both protein and peptide were binding competent, where a stoichiometry of $N = 1.250 \pm 0.004$ was observed, indicating that protein and peptide reagents were not the reason for failed SPR experiments. Allowing for slight errors in the calculation of peptide and protein concentrations, these results are also in accordance with the previously published data [102]. The initial project goals were to explore several pertubations of the Tankyrase binding peptides and thus experimental throughput was critical. While ITC is considered a gold-standard technique for measuring protein-ligand interactions, it is very expensive with regard to time and sample [217]. Thus we wanted to explore a more high-throughput approach, which could be used to screen several compounds in a much shorter time, orthogonally to FP experiments. As a proof of principle, we approached DSF with pep1 as a benchmark for the technique.

Figure 3.2 shows that TNKS1 ARC2-3 displays the characteristic DSF melting curve, with a linear pre-transition baseline, defined melt transition and linear post-transition baseline. As result of this typical behaviour, the model used to describe the data and the unfolding transition (see Equation 3.1), provides an accurate melting temperature for unfolding transition of TNKS1 ARC2-3. However, this is not seen to the same extent in TNKS2 ARC4 DSF experiments (Figure 3.2b). Instead, TNKS2 ARC4 shows a temperature dependant increase of the fluorescence between 20°C to 40°C that is does not adopt

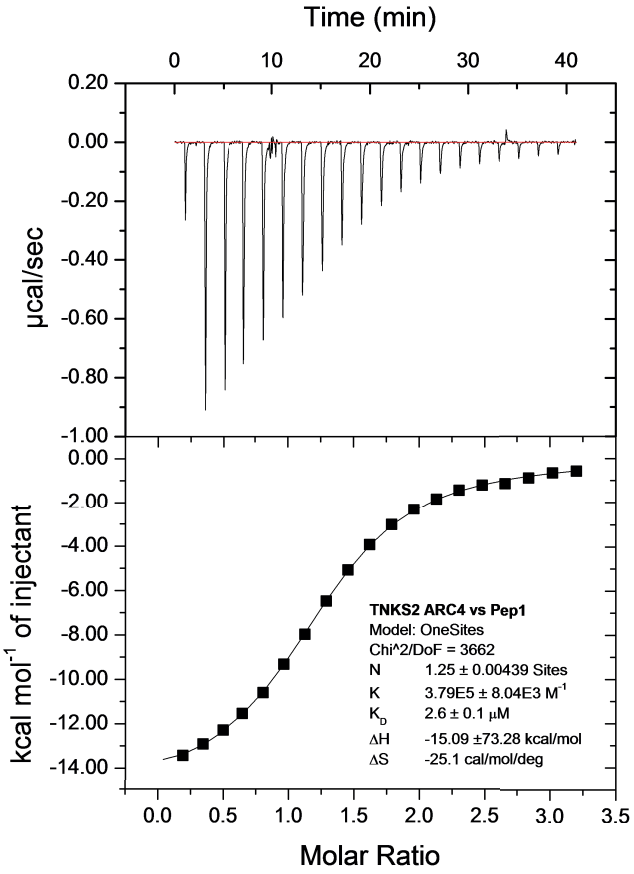
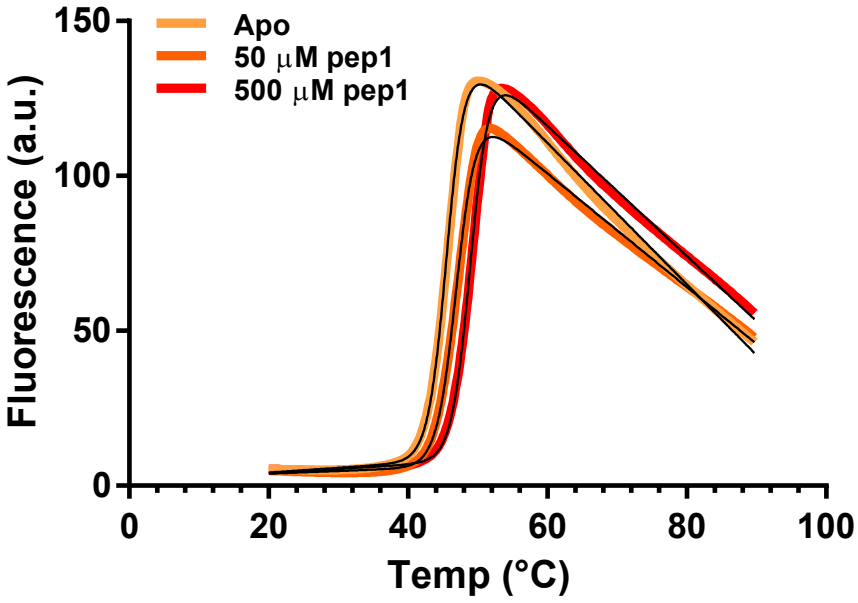
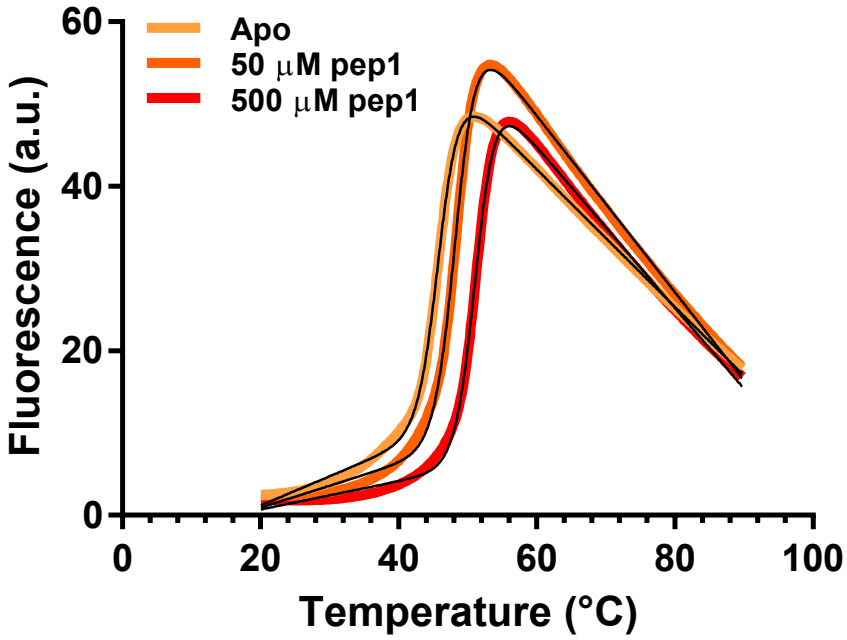


Figure 3.1: ITC thermogram of 300 µM pep1 against 20 µM TNKS2 ARC4 at 25°C in 50 mM Tris-HCl, 150 mM NaCl, 2 mM TCEP, pH 7.5.

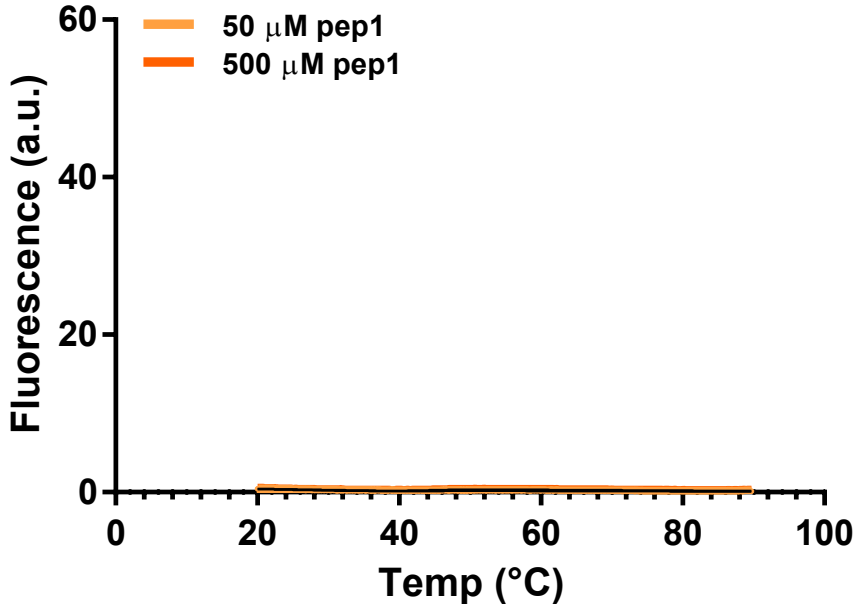
the linear nature that is described by our model, describing a two-state transition. This indicates that an intermediate state or more specifically that Sypro Orange dye is gradually binding and increasing its fluorescence intensity within this region. Nevertheless, there is a defined unfolding transition between 40°C and 60°C. Since the model doesn't describe this data properly in this early region, it is possible that absolute melting temperature calculations are systematically incorrect. The data for pep1 without any protein clearly indicate that the change in fluorescence with increasing temperature is due to TNKS unfolding. pep1 thermally stabilises TNKS2 ARC4 more than TNKS1 ARC2-3, at both 50 µM and 500 µM.



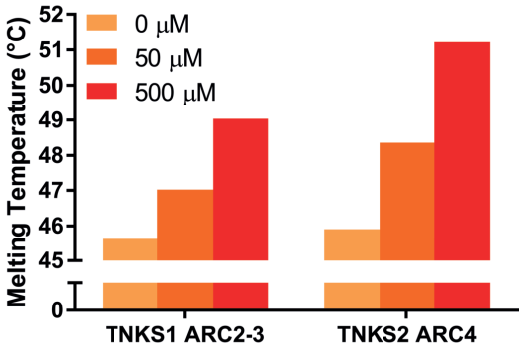
(a) TNKS1 ARC2-3



(b) TNKS2 ARC4

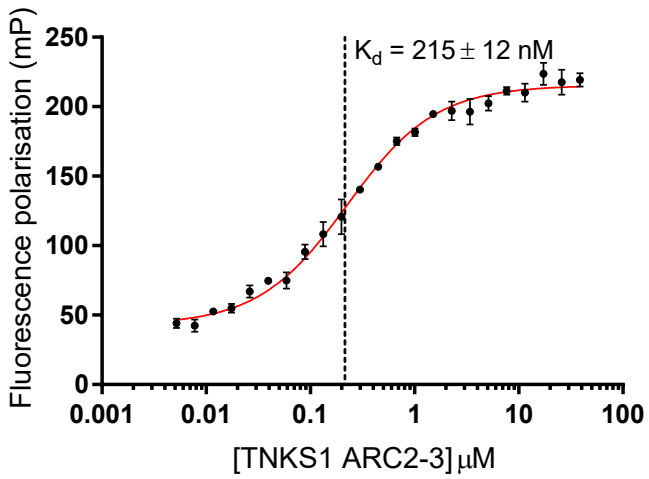


(c) Peptide only controls

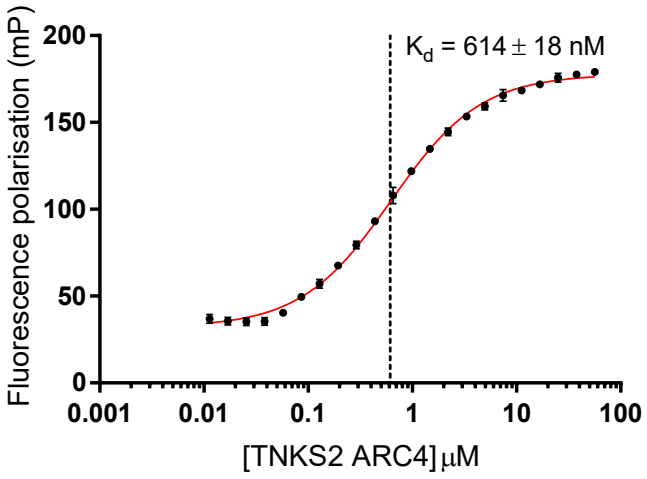


(d) T_m plots

Figure 3.2: DSF melt curves of (a) TNKS1 ARC2-3 and (b) TNKS2 ARC4 with or without pep1 at the indicated concentrations. (c) DSF melt curves of pep1 at the indicated concentrations without any TNKS proteins. (d) Calculated melting temperatures of TNKS proteins with and without pep1 at indicated concentrations. Black lines indicate the data fit using equation 3.1. Temperature ramp from 20°C to 90°C in 50 mM Tris-HCl, 150 mM NaCl, 2 mM TCEP, pH 7.5.



(a)



(b)

Figure 3.3: Titrations of (a) TNKS1 ARC2-3, and (b) TNKS2 ARC4 against 20 nM TAMRA-pep1. All reagents were diluted in PBS + 0.01% (v/v) Tween 20. Measurements read at 25°C. Black dashed lines indicate the calculated K_d of TAMRA-pep1 to the Tankyrase subdomains according to Equation 3.3.

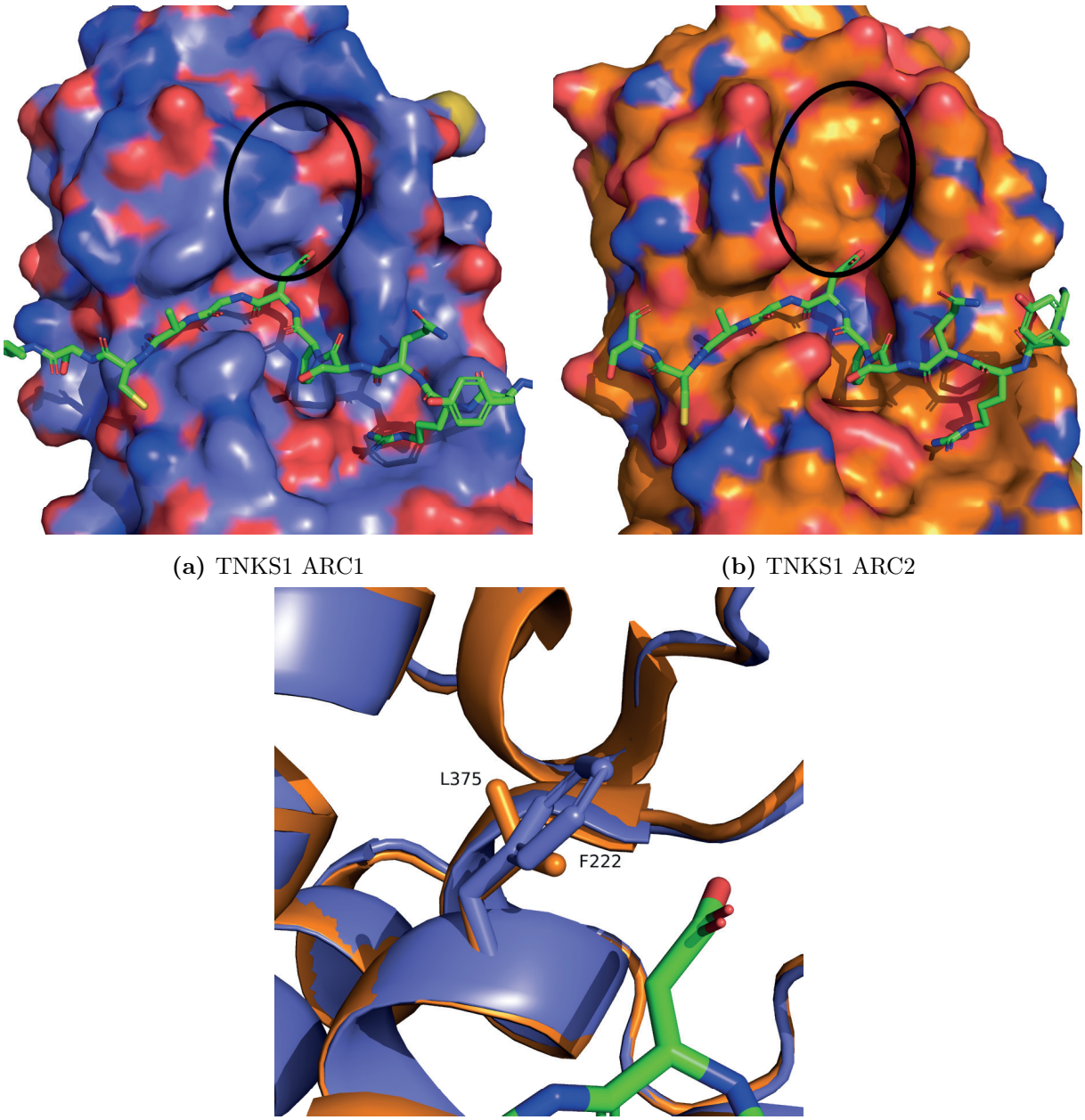
We then sought to improve the FP protocol described by Xu *et al.* [102] by transferring the assay into 384-well plate format. This allowed us to reduce sample volumes per well, while at the same time increasing the density of data points for a given end-point equilibrium titration experiment using a 1.5-fold serial dilution series over twenty points versus a 2-fold dilution series over twelve points. With the modified assay set up, we then confirmed the dissociation constant of the tracer peptide, TAMRA-pep1 against both TNKS1 ARC2-3 and TNKS2 ARC4 at 215 ± 12 nM and 614 ± 18 nM, respectively. This data allowed us to set up competition FP experiments to analyse pertubations of peptide inhibitors against these binding sites.

3.3.3 Towards selective inhibitory peptides of ARCs 2/5 versus ARCs 1/4

Despite the high sequence identity between the ARC domains of TNKS1 and TNKS2, there are structural differences around the peptide-binding site. Namely, in both TNKS proteins, ARCs 1 and 4 and ARCs 2 and 5 are most similar. Initially, we wanted to investigate whether we could exploit these differences highlighted in Figure 3.4, where a large and extended surface is potentially accessible in ARCs 2 and 5. In comparison, this binding surface is inaccessible in ARCs 1 and 4, due to the protrusion of the benzyl ring from F222 (ARC1) from the surface of the protein. In contrast, ARC2 has its equivalent residue in the ARC replaced by a smaller hydrophobic residue, L375. We believed we could make use of this extended surface on ARC2 and ARC5 to improve the affinity.

Since Guettler *et al.* only explored pertubations of Tankyrase binding peptides against the TNKS2 ARC4 domain [131], we hypothesised that those trends may not be entirely applicable to the ARC2 or ARC5 domains, given their structural differences. Although a pairwise matrix screen could have been performed against the TNKS1 ARC2-3 domain, it was believed that this would largely not result in a significantly improved compared to the pep1 sequence (Ac-REAGDGEE-NH_2) and indeed purchasing such a large set of peptides would be not be economical. To explore this hypothesis in a more rational and focused manner, a set of seven peptides were purchased from Mimotopes, to explore a series of natural amino acid point mutations at position 5 of the peptide, which are listed in Table 3.2. These peptides were specifically designed in an attempt to explore the region highlighted in Figure 3.4b, with a variety of the larger side chain-containing amino acids to explore various chemical properties associated with them, such as aromaticity, aliphatic characteristic and positive charge. Peptides T5-T11 were then tested in the competition FP assay to assess their dissociation constants to both the TNKS1 ARC2-3 and the TNKS2 ARC4 constructs.

It was evident from the data presented in Figure 3.5 that none of the purchased peptides in Table 3.2 were of similarly high affinity compared to pep1. At maximum concentrations of 1 mM, none were able to fully compete off the TAMRA-pep1 peptide in the context of TNKS1 ARC2-3 or TNKS2 ARC4. Due to this, any affinities calculated by Equation 3.4, will not be accurate and are therefore absolute affinities are not worth presenting here. Upon closer inspection of the data, we noticed that peptide T5 (Ac-REAGRGEE-NH_2) was able to displace TAMRA-pep in the context of TNKS1 ARC2-3, but not TNKS2 ARC4 (Figure 3.5c). These data suggests an apparent affinity that exists for TNKS1 ARC2-3, containing the Leucine residues versus the Phenylalanine residue



(a) TNKS1 ARC1

(b) TNKS1 ARC2

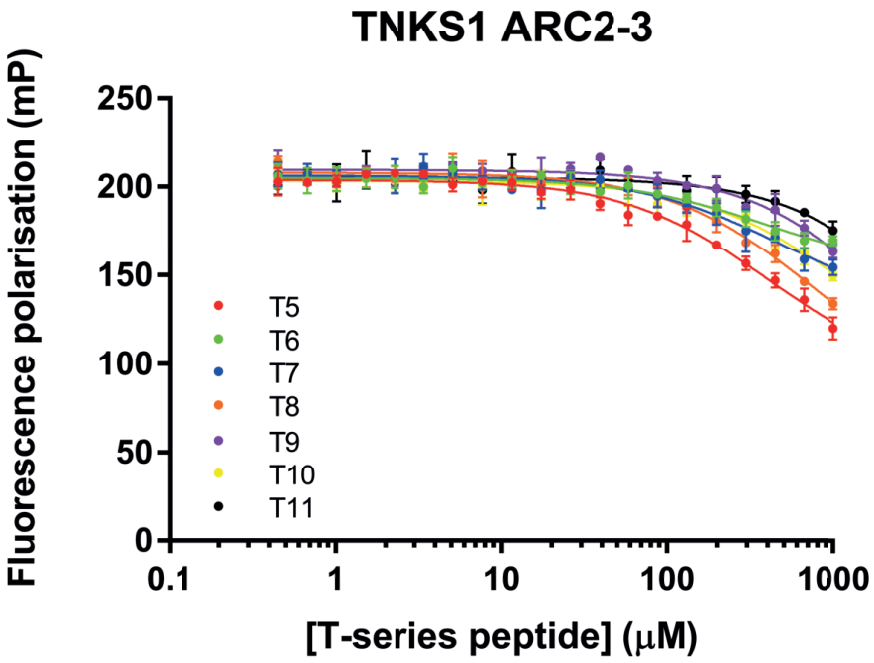
(c) ARC1 - ARC2 alignment

Figure 3.4: IRAP peptides bound to (a) Human TNKS1 ARC1 and (b) Human TNKS1 ARC2. Black circles indicate area of interest to exploit for further rational design. (c) Structural alignment of human TNKS1 ARC1 and ARC2. F222 and L375 are highlighted. Figures produced from PDB ID: 5JHQ [126].

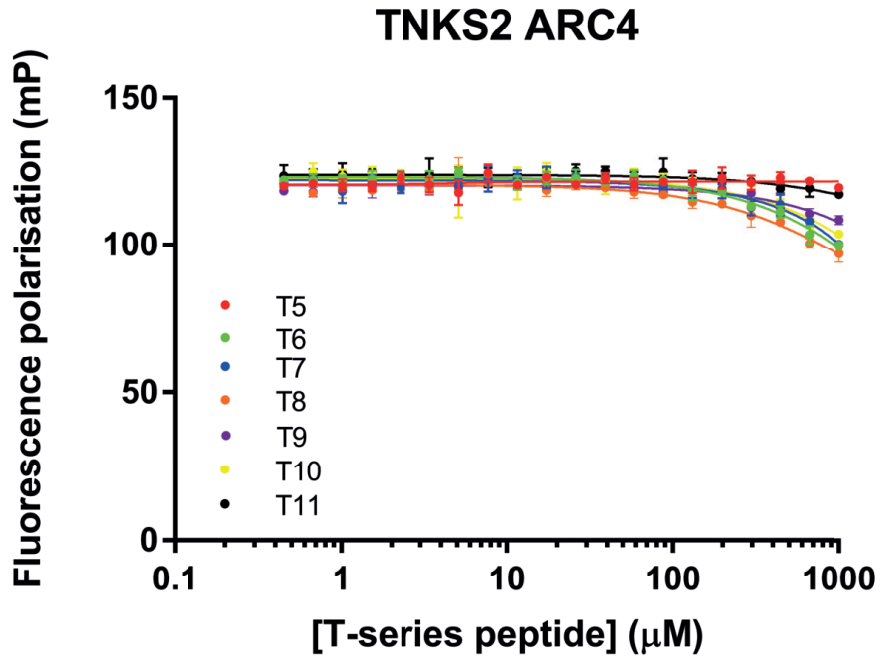
found in the ARC1 and ARC4 domains. This indicates that this surface may be conducive for further inhibitor design efforts to specifically target the ARC2 and ARC5 domains of Tankyrases.

Table 3.2: Peptides exploring natural amino acid point mutations at position 5 (highlighted in red), with a bias towards large side chains. All purchased peptides were >95% purity by analytical HPLC.

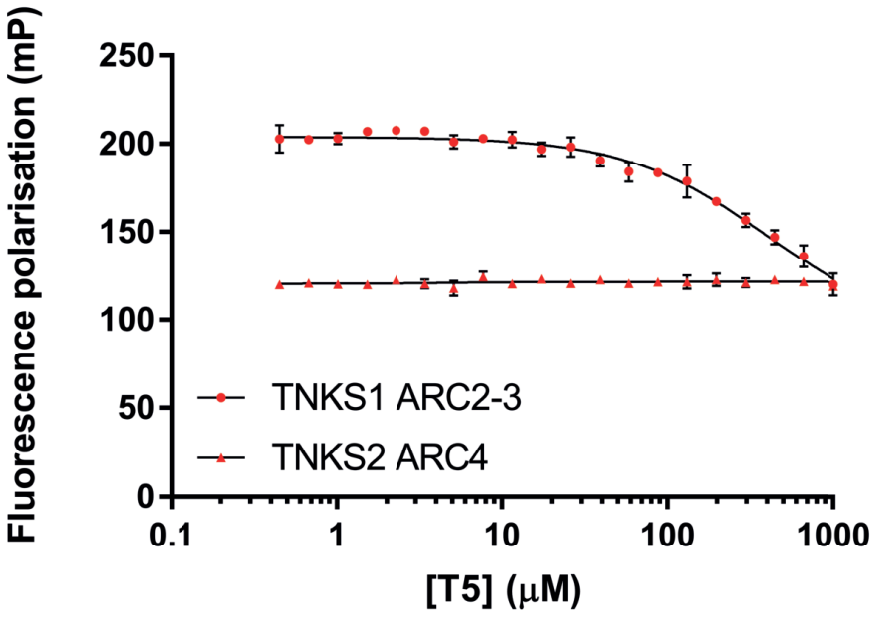
Peptide	Sequence
T5	Ac-REAGRGEE-NH_2
T6	Ac-REAGHGEE-NH_2
T7	Ac-REAGLGEE-NH_2
T8	Ac-REAGMGEE-NH_2
T9	Ac-REAGFGEE-NH_2
T10	Ac-REAGYGEE-NH_2
T11	Ac-REAGWGEE-NH_2



(a)



(b)



(c)

Figure 3.5: Competition FP experiments with T-series peptides against (a) TNKS1 ARC2-3, TAMRA-pep1 complex and (b) TNKS2 ARC4, TAMRA-pep1 complex. (c) Competition FP experiments, with T5 competing for TNKS1 ARC2-3 (red circles) or TNKS2 ARC4 (red triangles), in the presence of TAMRA-pep1. Protein and peptide solutions were diluted in PBS+0.01% Tween 20, 1% (v/v) DMSO. All measurements read at 25°C.

3.3.4 Macrocycle interactions to TNKS

Peptide design

Since the previously described efforts did not lead to any improved binding affinity to the TNKS proteins, we next took some inspiration from the literature. It has long been known that macrocycles in stapled peptide inhibitors are able to improve their affinity to the target protein by clustering peptide conformations towards the bound state of the ligand, thus reducing the entropic cost of binding. One of the more studied stapled peptide systems, targeting the p53-Mdm2 interaction, has garnered a lot of attention over the years from both academia and the pharmaceutical industry. One such inhibitor of Mdm2, peptide M06, was observed to make use of the macrocycle constraint to produce favourable interactions with the surface of Mdm2 as depicted in Figure 3.6 [218]. We attempted to utilise this principle in the design of our macrocyclised peptides against Tankyrase.

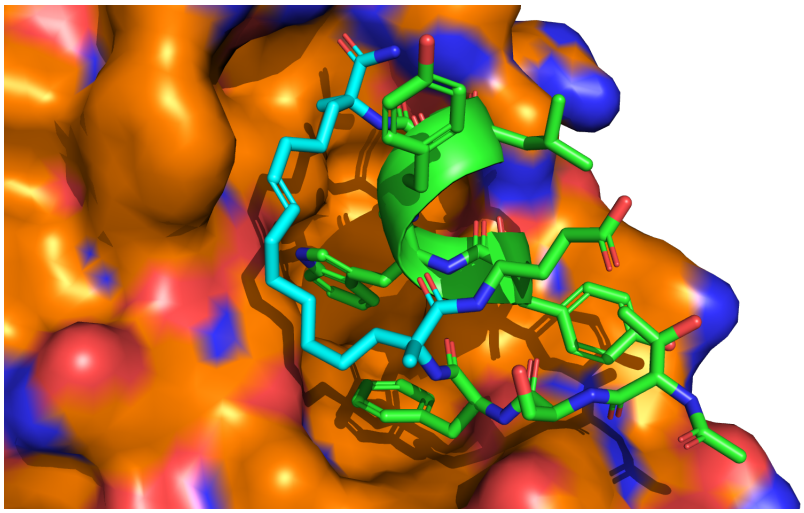
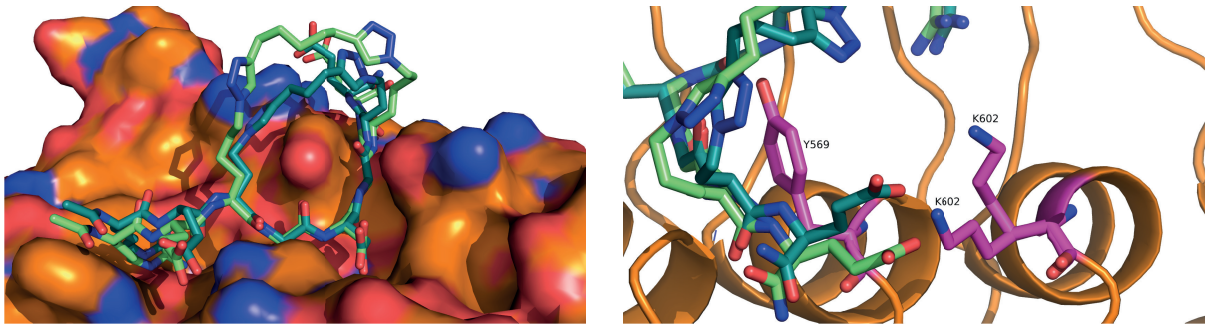


Figure 3.6: X-ray crystal structure of M06 stapled peptide bound to Mdm2. Peptide (green) is depicted in cartoon and stick configuration, with staple amino acids coloured separately in cyan. Mdm2 surface is shown in orange. PDB ID: 4UMN [218].

To approach this, we first analysed the co-crystal structures of TNKS2 ARC4 bound to sp4n2m3 (PDB ID: 5BXO) and sp4n4m5 (PDB ID: 5BXU) [102]. In both structures, the macrocycle is inserted at positions 3 and 7 in the primary amino acid sequence. While both peptides were slightly improved in their affinity to TNKS2 ARC4, the macrocycles in both cases were solvent exposed and thus not able to contribute an interaction with the protein, (Figure 3.7a). We then decided to change the macrocycle positioning in an attempt to dock the macrocycle within the groove formed between Y569 and K602, (Figure

3.7b). Although this groove is not particularly well defined in the crystal structure, K602 is resolved in two different conformations implying a degree of flexibility in the side chain that could perhaps be exploited for our design.



(a) **(b)**

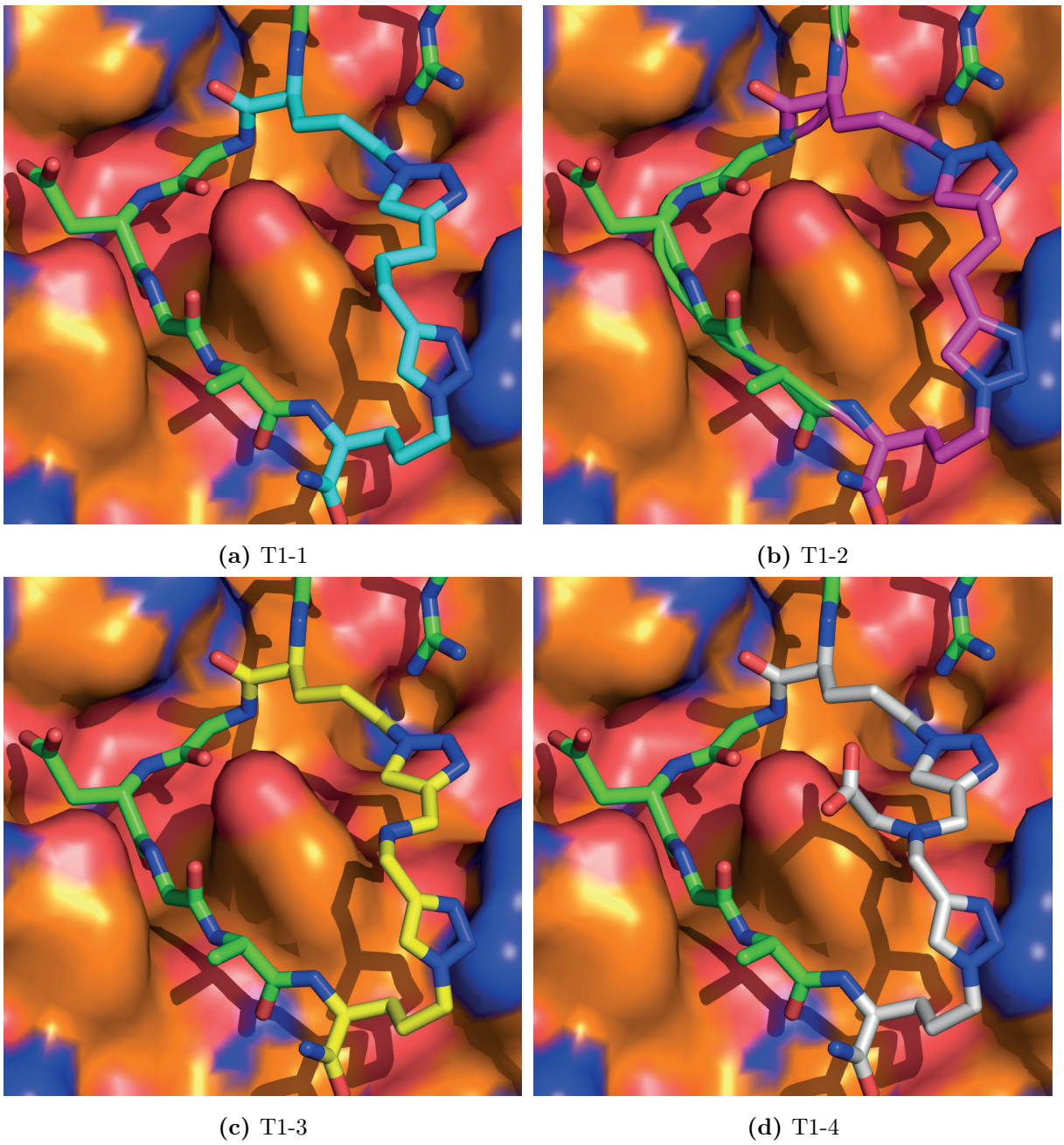
Figure 3.7: (a) Alignment of sp4n2m3 (turqoise) and sp4n4m5 (pale green) peptides bound to TNKS2 ARC4. (b) Groove formed between Y569 and K602 between the ANK repeats.

With guidance from our collaborator, Dr. Yaw Sing Tan (BII, A*STAR, Singapore), we re-designed the macrocycle to positions 3 and 8 in the primary sequence. We believed this may work since the side chain from Glu8 already points towards the defined groove (Figure 3.7b). This led to the design of the T1 series of peptides using L-Ornithine (N_3) unnatural amino acids, and differing dialkynyl linkers. Their proposed binding models are depicted in Figure 3.8. At the time of design, we favoured the hepta-1,6,-diyne linker due to the availability of similar linkers to which additional modifications could be incorporated downstream, such as the addition of a cell-penetrating peptide to the carboxyl group of the T1-4 linker.

The synthesis of T1 and subsequently T1-1 was relatively simple and straightforward. However, practically it was fraught with difficulty due to the hygroscopic properties of the peptides. This was made more difficult by the low milligram scale at which synthesis, purification and double-click reactions were performed. For this reason, the other iterations of T1 peptides were not synthesised. Peptide synthesis data is provided in Appendix B.

Biophysical characterisation of T1-1

Due to the hygroscopic nature of T1-1, we could not be confident in the concentration of dissolved peptide from the mass of peptide weighed out. A small sample was submitted for amino acid analysis (Dr. Peter Sharrat, Former PNAC facility, Department of Biochemistry, University of Cambridge), where the concentration of the working stock solution was confirmed. We then determined the dissociation constant of T1-1 by ITC



(a) T1-1 **(b)** T1-2

(c) T1-3 **(d)** T1-4

Figure 3.8: *In silico* models of T1 series peptides bound to TNKS2 ARC4. (a) T1-1. Linker: hepta-1,6-diyne. (b) T1-2. Linker: hexa-1,5-diyne. (c) T1-3. Linker: di(prop-2-yn-1-yl)amine (d) T1-4. Linker: di(prop-2-yn-1-yl)glycine.

and competition FP, where we observed binding by both techniques. Using ITC we determined the peptide dissociation constant of 6.62 ± 0.13 μM, which was 2.5-fold weaker than pep1 measured by ITC (Figure 3.1). Likewise, using competition FP we observed a similar trend, obtaining a dissociation constant of 3.41 ± 0.08 μM, which was 5.5-fold weaker than TAMRA-pep1 measured by direct FP (Figure 3.3b). Co-crystallisation of the peptide was then attempted with TNKS2 ARC4. However, despite numerous attempts following protocols/ conditions described by Xu *et al.* [102], none were successful in pro-

ducing crystalline material. Lastly, Apo-TNKS2 ARC4 crystals were produced and it was attempted to obtain co-crystals by ligand-soaking. However, these approaches were also unsuccessful.

3.3.5 Exploring Arginine mimetics

Our final attempt at improving the binding affinity of TNKS ARC-binding peptides was also guided by structural information. This effort focused around the conserved arginine residue at position 1 of the peptide sequence. Guettler *et al.* had described this amino acid as being crucial for binding, adopting the 'arginine cradle' site (Figure 3.10) [131]. Analysing several peptide bound structures, the backbone atoms of residues 1-3 of the peptide pull away from the protein surface in order to accommodate the large arginine sidechain. Since it is understood that this interaction between the guanidinium moiety of the arginine sidechain is necessary, we explored synthesising a TNKS peptide with a shorter arginine sidechain. Specifically, shortening the chain length by one or two

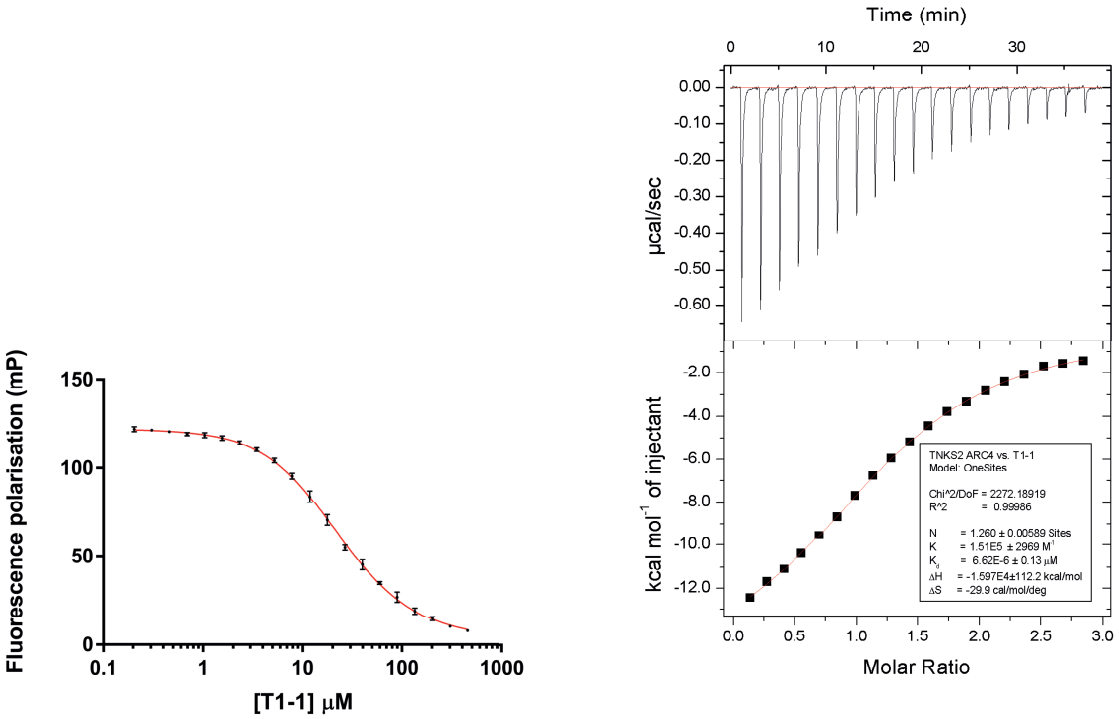


Figure 3.9: Biophysical characterisation of peptide T1-1 binding to TNKS2 ARC4. (Left) Competition FP experiment titrating T1-1 against a TNKS2 ARC4, TAMRA-pep1 complex. Protein and peptide solutions were diluted in PBS+0.01% Tween 20, 1% (v/v) DMSO. All measurements read at 25°C. (Right) ITC thermograms of T1-1 titrated into the cell containing TNKS2 ARC4 at 25°C in 50 mM Tris-HCl, 150 mM NaCl, 2 mM TCEP, 1% (v/v) DMSO, pH 7.5.

methylene groups may enable the backbone atoms of residues 1-3 to form interactions with TNKS2 ARC4.

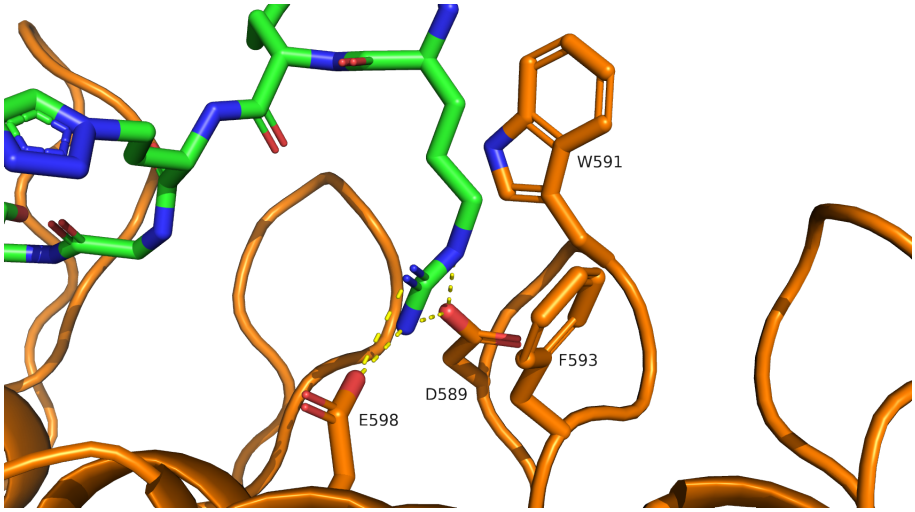


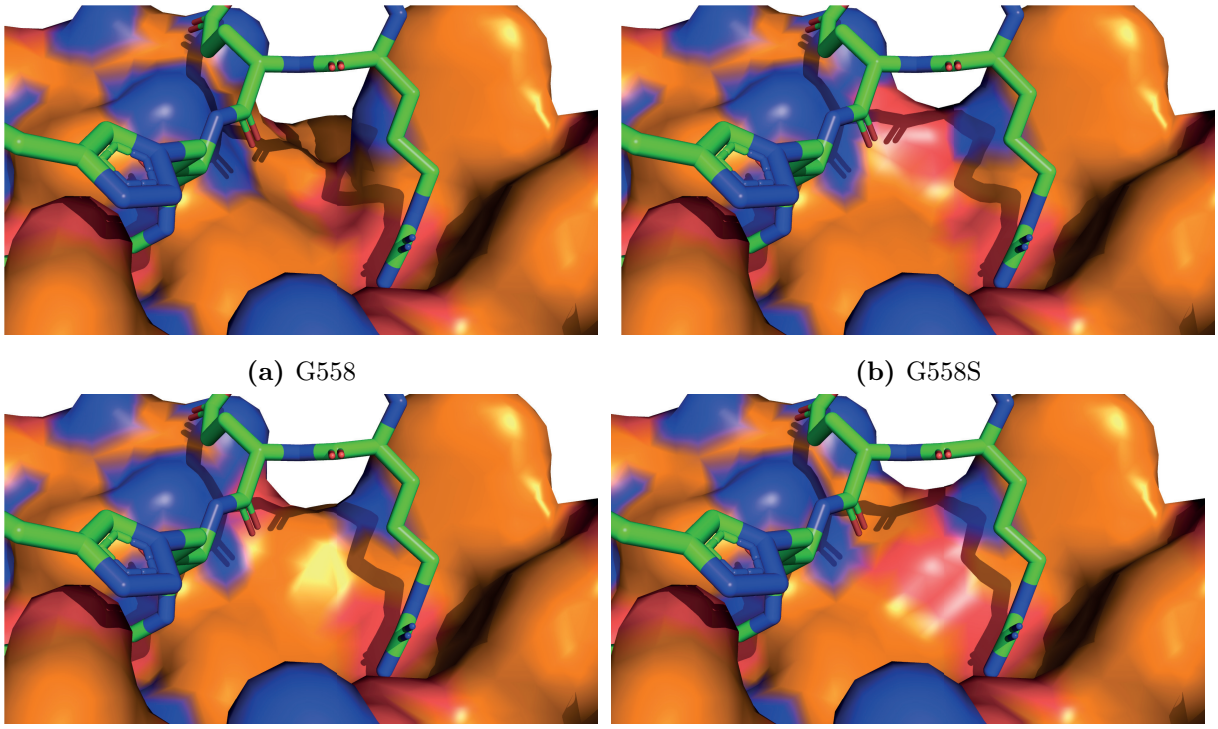
Figure 3.10: Schematic of arginine 1 of TNKS binding peptides binding in the 'arginine cradle' site, composed of D589, W591, F593 and E598 of TNKS2 ARC4.

Unfortunately, obtaining these modified Fmoc-amino acids or similar derivatives was prohibitively expensive, thus we attempted to approach this from the opposite perspective. We sought to mutate a non-conserved residue of the ANK repeat, G558 of TNKS2 ARC4 in order to ascertain whether it was possible to gain affinity through extra interactions to the peptide. G558 is located in the loop region of the ANK repeat prior to the repeat containg W591. We visualised these mutations by simple *in silico* mutagensis of the 5BXO structure in Figure 3.11 (a-d).

We then assayed wild-type GST-TNKS2 ARC4 and its mutants in a direct FP assay against TAMRA-pep1 (Figure 3.12 and Table 3.3), to see whether we had improved the binding affinity by introducing additional interactions to the peptide. Unexpectedly, the hypothesis proved incorrect, with a loss in binding affinity observed across all three mutants. The decrease in affinity appeared to correlate with an increase in side chain length. However, it was unclear whether this effect was due to the mutation disrupting the structure of the protein, or whether these residues directly disrupted the interaction.

3.4 Discussion

We first aimed to expand the repertoire of biophysical assays to assess ligand binding using SPR. Although binding was observed, the response of a known peptide analyte was significantly lower than expected. It is known that some proteins do not tolerate the immobilisation process often used in SPR, whereby proteins are covalently coupled to



(a) G558 **(b)** G558S

(c) G558T **(d)** G558D

Figure 3.11: *In silico* mutagensis models of TNKS2 ARC4 mutations adjacent to the 'arginine cradle' site.

Table 3.3: Dissociation constants of TAMRA-pep1 against GST-TNKS2 ARC4 (G558X) mutants.

Protein	Dissociation constant, K_d (nM)	Fold change
G558	320 ± 14	-
G558S	1229 ± 55	4
G558T	987 ± 32	3
G558D	4410 ± 226	14

the dextran matrix by EDC/NHS coupling. Furthermore, cross-linking can occur near the analyte-binding site and thus occlude the analyte sterically. In addition, EDC/NHS coupling is not performed in a site-specific manner, such that multiple surface-exposed primary amines can be modified within the same protein molecule. According to a study by Papalia and Myszka, a coupling stoichiometry of the protein to the surface that is greater than 1:1 is detrimental to the activity of the ligand [219]. Unfortunately we only became aware of this work after the Tankyrase project had been discontinued. Otherwise, it would have been beneficial to explore non-covalent attachment strategies to establish a more robust TNKS SPR assay.

In parallel to SPR experiments, binding affinites were also assessed using two orthog-

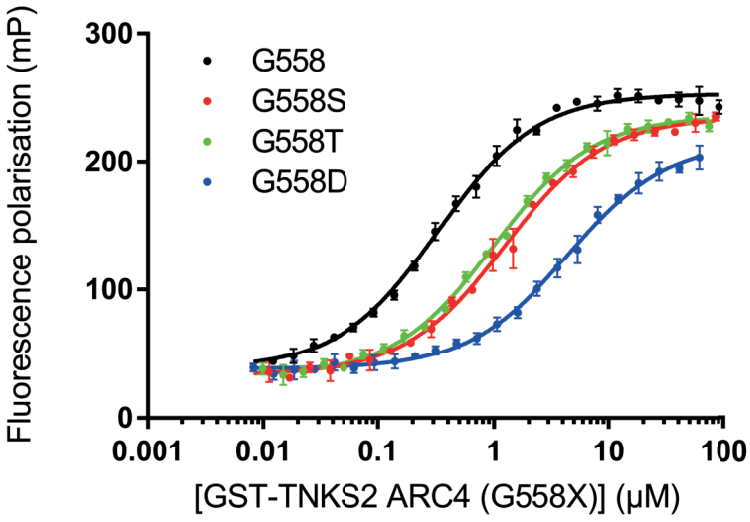


Figure 3.12: Direct FP measurements of TAMRA-pep1 against wild-type GST-TNKS2 ARC4 (G558) and mutants G558S, G558T and G558D. Protein and peptides were diluted in PBS+0.01% (v/v) Tween 20 and measurements read at 25°C.

onal methods, DSF and ITC. Since some ITC data had already been published, it was clear that the assay would work as expected. The data subsequently obtained confirmed that the poor data from SPR experiments was unlikely due to the issues with protein or peptide samples, such as aggregation or partial unfolding. However, for the purpose of ligand screening, ITC is a low throughput method and therefore, DSF was explored to overcome this limitation. Initial experiments were conducted in a 384-well plate format using a Roche LightCycler 480 II at AstraZeneca. However this instrument was not routinely available and hence subsequent DSF experiments were performed in a 96-well plate format using a Biorad CFX Connect machine located in another department of the university. These two instruments were found to differ greatly in the quality of raw data produced. Since the data obtained from the Biorad machine was of insufficient quality to be further analysed, the DSF data presented here was limited to the initial experiments performed with pep1 at AstraZeneca. All melt curves generated with TNKS1 ARC2-3 could be fitted using Equation 3.1 resulting in accurate melting temperatures, whereas transition mid-points of TNKS2 ARC4 melt curves were skewed by the non-linear pre-transition behaviour.

The next goal was to improve the affinity of peptide ligands targeting the ARC domains. Instead of following the same strategy as set out by Dr. Wenshu Xu, it was thought most reasonable to improve the peptide prior to macrocyclisation, for example by developing ligands that could distinguish between the binding surface of TNKS1 ARC2-3 and TNKS2 ARC4. While the peptides listed in Table 3.2 did not bind with high affinity to either TNKS construct, ARC specificity was observed peptide T5. ARC selectivity may

be relevant in the regulation of Axin1 as Eisemann *et al.* observed wild-type Axin 1 to bind primarily to ARC2 and ARC5 simultaneously rather than ARC1 and ARC4 [126]. Since Tankyrases rely on avidity [129], the rationale to selectively inhibit the ARC domains responsible for Axin1 binding may be beneficial. Although the data presented here suggests that such selectivity can be achieved, at the time it was not possible to further the T5 peptide as it required medicinal chemistry knowledge.

Given the limited success of improving the natural peptide sequence, macrocycles were re-designed with the aim of introducing the macrocycle-protein interactions through hydrogen-bonding. One peptide, T1-1, was successfully synthesised and characterised although its binding affinity to TNKS2 ARC4 was at least 2-fold weaker than pep1. Co-crystallisation of the TNKS2 ARC4-T1-1 peptide complex were attempted to further understand why the peptide is weaker and where subsequent modification could be introduced to improve the it. However, despite multiple attempts all efforts were unsuccessful. Given the practical issues involved with this particular peptides due to its severely hygroscopic nature, we ultimately ceased to pursue this approach to the project.

3.5 Conclusions

The project to inhibit the Tankyrase substrate binding domains continued on from a project of a previous PhD student. The project aimed to further improve the peptide binding affinity to the ARC domains towards that of a typical therapeutic molecule. This was a significant challenge since work by Guettler *et al.* and Dr. Wenshu Xu had seemingly explored these peptides in a thorough manner. Nonetheless, it was possible to identify a peptide that displayed selectivity towards TNKS1 ARC2-3 over TNKS2 ARC4, in addition to rationally design novel macrocycles. Lastly, TNKS2 ARC4 G558X mutants were used to explore the surface cavity adjacent to the arginine cradle site, which indicated that G558 was either necessary for binding or crucial to protein stability.

Chapter 4

Production of Cdc20 for biophysical characterisation

4.1 Introduction

The laboratories of David Barford and Brenda Schulman have published the expression and purification of full-length human Cdc20. These groups have primarily used full-length Cdc20 to obtain cryogenic electron microscopy (cryo-EM) structures of the APC/C in the context of its co-activators, substrates and sub-complexes [28, 150–154]. However, the APC/C co-activators, Cdc20 and its homolog Cdh1, contain intrinsically disordered regions (IDRs) at the N- and C- termini (Figure 4.1). These IDRs are not responsible for binding to the APC/C substrates and were considered unnecessary for the biophysical characterisation of designed peptide molecules. The laboratory of Xuelian Luo has published multiple X-ray crystal structures of the $Cdc20^{WD40}$ domain, containing residues 161-477 [193, 195]. Since one of the project goals was to produce co-crystal structures of Cdc20 bound to inhibitory peptides, it was rational to focus construct design around published data. Over the years, all groups producing recombinant Cdc20 have used baculovirus expression vector systems (BEVS). There are no reports on the expression of Cdc20 using *E. coli*, and conversations with David Barford highlighted the necessity of eukaryotic chaperones to aid the folding of Cdc20. Therefore, BEVS were used to generate Cdc20 protein for biophysical characterisation.

4.2 Methods

4.2.1 Molecular biology

A plasmid encoding His_6-MBP-$Cdc20^{1-499}$ with a TEV cleavage site between the MBP and Cdc20 genes in the pU1 vector was a kind gift from Ziguo Zhang and David Barford

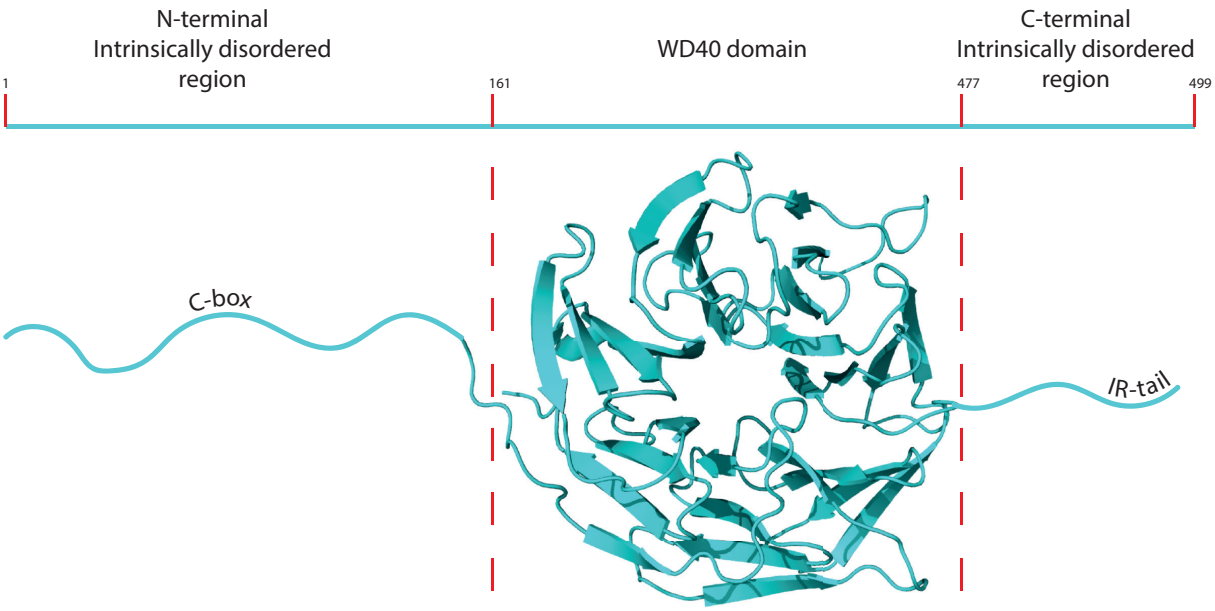


Figure 4.1: Illustration of Cdc20 domains. The WD40 domain is flanked by two intrinsically disordered regions. The N-terminal region spans residues 1-160, containing the C-box motif. The WD40 domain spans residues 161-477 and is responsible for binding to substrate proteins via degron motifs. The C-terminal intrinsically disordered region comprises residues 478-499, with the IR-tail comprising the terminal residues. The WD40 domain depicted is from PDB: 4GGC, [193].

(MRC-LMB, Cambridge) [200].

RTH mutagenesis (see Section 2.2.2) was used to remove the MBP tag and the N- and C- terminal IDRs of Cdc20 in an iterative process, creating a variety of intermediate constructs that could be useful for this project or for the Barford laboratory (Table 4.1). In addition to producing recombinant Cdc20 in BEVS, full-length Cdc20 was tagged with the HiBiT peptide (Promega) in a pcDNA 3.1(-) vector for transient transfection in mammalian cells. All plasmids and bacmids were produced as described in Section 2.2.6 and are listed in Table 4.1.

4.2.2 Baculovirus expression of recombinant Cdc20 constructs

Generating recombinant virus

Insect cells were transfected in duplicate with bacmid DNA in 24 deep-well blocks. First, 20 µL of SuperFect transfection reagent (Qiagen) and 10 µg of Bacmid DNA (1µg/µL) were added to 150 µL of SF900 II medium (Gibco) in each well and were incubated at room temperature for 30 minutes to allow the Bacmid-SuperFect complex to form. To each well, 5 mL of mid-log Sf21 cells at 5×10^5 cells/mL in Sf900 II medium, supplemented with 50 µg/mL gentamicin and 10 % (v/v) FCS, were added. Blocks were sealed with

Table 4.1: Cdc20 constructs produced for BEVS or transient transfection in mammalian cell lines. Theoretical molecular weights (MW) and extinction coefficients at 280 nm (ϵ_{280}) are listed for the total gene product and for the final desired product after cleavage of the purification tag. Molecular weights and extinction coefficients were calculated using the Expasy ProtParam Web-tool [202].

Construct	Total gene product		Cdc20 protein	
	MW (Da)	ϵ_{280} ($M^{-1}cm^{-1}$)	MW (Da)	ϵ_{280} ($M^{-1}cm^{-1}$)
His_6-MBP-TEV-Cdc20 (1-499)	99526.0	178190	54850.7	105880
His_6-MBP-TEV-Cdc20 (1-477)	96926.0	178190	52250.7	105880
His_6-MBP-TEV-Cdc20 (161-499)	82122.4	162720	3447.1	90410
His_6-MBP-TEV-Cdc20 (161-477)	79522.4	162720	34847.2	90410
His_6-MBP-TEV-Cdc20 (165-477)	79049.8	162720	34374.6	90410
His_6-TEV-Cdc20 (161-477)	39331.7	96370	34847.2	90410
His_6-TEV-Cdc20 (165-477)	38859.2	96370	34374.6	90410
HiBiT-Cdc20 (1-499)	56444.6	-	54722.6	-
Cdc20 (1-499)-HiBiT	56588.8	-	54997.9	-

a breathable membrane and incubated at 27°C, shaking at 700 rpm (small orbit) for 7 days. After the incubation period, cells were pelleted at 3400×g, 4°C for 10 minutes and supernatant containing viral stocks of duplicate wells were pooled and stored at 4°C.

Baculovirus test expressions

Insect cell small-scale expression tests were performed in two formats which were dependant on the location where the test expressions were performed (AstraZeneca or Department of Pharmacology).

At AstraZeneca, mid-log Sf21 cells were added to a 24 deep-well block at 3×10^5 cells/mL with 180 µL of the relevant virus added. In addition to the various Cdc20 con-

structs, a virus expressing Yellow Fluorescent Protein (YFP) was included as a positive control. Blocks were sealed with a breathable membrane and incubated at 27°C, shaking at 700 rpm (small orbit) for 48 hours. Following the incubation period, cells were counted using a Vi-Cell cell counter (Beckman Coulter) prior to harvesting. The cell suspension was pelleted at 3400×g, 15 minutes, 4°C. The resulting pellet was resuspended in 1 mL PhyNexus wash buffer (40 mM HEPES, 300 mM NaCl, 20 mM imidazole, 1 mM TCEP, 1 × cOmplete^{TM} mini EDTA-free protease inhibitor tablet (Roche), 1 × benzonase, pH 7.4). Resuspended cells were freeze-thaw lysed and centrifuged as above to separate cellular debris. Supernatants were loaded onto a PhyNexus robot with Immobilised Metal ion Affinity Chromatography (IMAC) tips (PhyNexus). IMAC resins were equilibrated in 1 mL wash buffer, supernatants loaded and then washed twice in 1 mL wash buffer, prior to recovering the protein in elution buffer (40 mM HEPES, 300 mM NaCl, 500 mM imidazole, 1 mM TCEP, pH 7.4). Elution fractions were analysed using a NuPAGE 4-12 % acrylamide gel (Novex).

At the Department of Pharmacology, mid-log Hi5 or Sf9 cells (25 mL) were added to a 50 mL erlenmeyer flask at 5×10^5 cells/mL with 100 µL of His_6-TEV-Cdc20 (161-477) virus. The cell cultures were incubated at 27°C, shaking 140 rpm for 48 hours. After incubation, cells numbers were determined using a LUNA-II cell counter (Logos Biosystems). The cells were then pelleted at 7000×g, 4°C for 5 minutes and resuspended in 10 mL IMAC wash buffer (50 mM Tris-HCl, 300 mM NaCl, 10 mM imidazole, 1 mM $MgCl_2$, 5 % (v/v) glycerol, 1 mM TCEP, 1 SigmaFAST EDTA-free protease inhibitor tablet, pH 8.5). Resuspended cells were freeze-thaw lysed in liquid N_2 and were centrifuged at 15,000×g, 4°C for 45 minutes. Clarified lysates were loaded onto a 1 mL HisTrap Excel column at 1 mL/min via a sample line on an ÄKTA Pure chromatography system. The column was washed with 10 CV of IMAC wash buffer before eluting in IMAC elution buffer (50 mM Tris-HCl, 300 mM NaCl, 300 mM imidazole, 1 mM $MgCl_2$, 5 % (v/v) glycerol, 1 mM TCEP, pH 8.5). Samples were prepared and loaded onto a 12 % (w/v) acrylamide gel and run as described in 2.3.1.

Large-scale expression cultures

Sf21 or Sf9 cells were used for large scale expression cultures and were grown in Sf900 II (Gibco) or Insect-XPRESS (Lonza) medium, respectively. To expand the virus stock, so-called 'short virus preps' were prepared, which were later added to the final expression culture volume. On day 1, virus was added (1 mL virus/ 500 mL cells) to cells at 1.5×10^6 cells/mL, generating the 'short virus prep'. At the same time, cell cultures were expanded to the desired culture volume, seeded at a density of 3×10^5 cells/mL. Cell cultures and the 'short virus prep' were incubated at 27°C, shaking at 140 rpm for 72 hours. The

'short virus prep' and cell cultures were counted before diluting the 'short virus prep' into the expanded cells (25 mL short virus prep/ 500 mL expanded cell culture). The cell cultures were incubated at 27°C, shaking at 140 rpm, for 48 (Sf21) or 60 (Sf9) hours. Final cell counts were measured and cells were harvested by centrifugation at 3400×g, 4°C, 15 minutes. Cell pellets were stored at -80°C until purifications were performed.

4.2.3 Purification of Cdc20^{WD40} protein

Purifications followed a general workflow as illustrated in Figure 4.2. The purification was guided by details contained in previous publications [193, 195]. Due to multiple issues during successive purifications, each attempt differed in its methodology to improve the yield and purity of Cdc20. Buffers used in purifications are listed in Table 4.2.

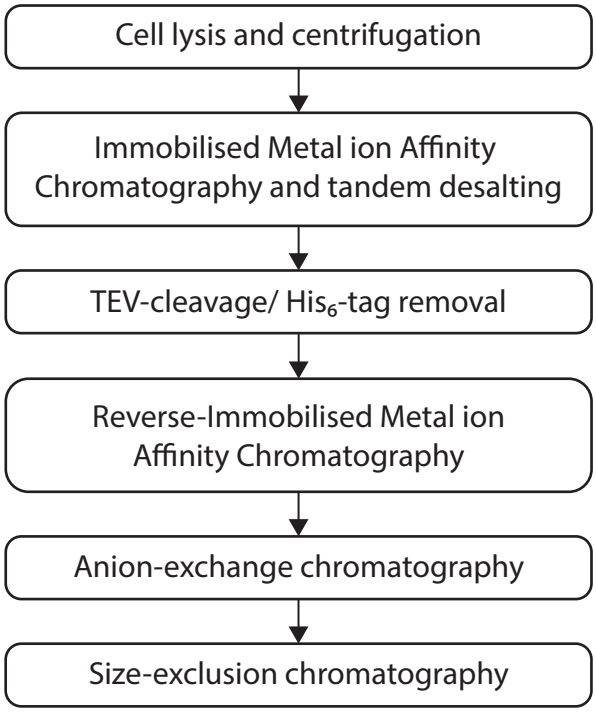


Figure 4.2: General workflow of Cdc20 purifications.

Cell lysis

Cell pellets (pellet from 1 L of cell culture) were resuspended in 100 mL IMAC wash buffer with SigmaFAST EDTA-free protease inhibitor tablets (Sigma) and crude DNaseI (Sigma) or Basemuncher (Expedeon), using a serological pipette or a magnetic stirrer until the lysate was homogenous. Lysis buffer volume scaled accordingly to the amount of cell pellet purified. Resuspended cells were transferred into a resealable plastic bag or aliquotted into 15 mL falcon tubes, frozen flat on dry ice or in liquid N_2, respectively, and

Table 4.2: Buffers used in Cdc20 purifications. Initial purifications began with first generation buffers and gradually transitioned to second generation buffers by successive optimisation.

Chromatography step/ buffer	First generation	Second generation
IMAC wash	50 mM Tris-HCl, 300 mM NaCl, 1 mM $MgCl_2$, 30 mM imidazole, 5% (v/v) glycerol, 0.25 mM TCEP, pH 7.5	50 mM Tris-HCl, 300 mM NaCl, 1 mM $MgCl_2$, 10 mM imidazole, 5% (v/v) glycerol, 1 mM TCEP, pH 8.5
IMAC elution	50 mM Tris-HCl, 300 mM NaCl, 1 mM $MgCl_2$, 300 mM imidazole, 5% (v/v) glycerol, 0.25 mM TCEP, pH 7.5	50 mM Tris-HCl, 300 mM NaCl, 1 mM $MgCl_2$, 300 mM imidazole, 5% (v/v) glycerol, 1 mM TCEP, pH 8.5
IEX LSB	25 mM Tris-HCl, 1 mM $MgCl_2$, 0.25 mM TCEP, pH 7.5	25 mM Tris-HCl, 1 mM $MgCl_2$, 5% (v/v) glycerol, 1 mM TCEP, pH 8.5
IEX HSB	25 mM Tris-HCl, 1 M NaCl, 1 mM $MgCl_2$, 0.25 mM TCEP, pH 7.5	25 mM Tris-HCl, 1 M NaCl, 1 mM $MgCl_2$, 5% (v/v) glycerol, 1 mM TCEP, pH 8.5
SEC	25 mM Tris-HCl, 150 mM NaCl, 1 mM $MgCl_2$, 5% (v/v) glycerol, 1 mM TCEP, pH 8.5	25 mM Tris-HCl, 150 mM NaCl, 1 mM $MgCl_2$, 5% (v/v) glycerol, 5 mM TCEP, pH 8.5

then thawed in a room temperature water bath. Lysates were clarified by centrifugation at 45,000×g, 4°C for 45 minutes. Supernatants were pooled for subsequent IMAC.

IMAC

Supernatants were loaded onto a 5 mL HisTrap Excel column equilibrated in IMAC wash buffer at 3 mL/min via the sample pump of an ÄKTA Pure chromatography system. The column was washed using IMAC wash buffer at 7.5 mL/min for 20-30 CV. Proteins were eluted as described in Section 2.3.2 into a 96 deep-well block, 50 mL falcon tubes (pre-filled with IMAC wash buffer, no imidazole) or via the outlet valve into a duran bottle (pre-filled with IMAC wash buffer, no imidazole) on a magnetic stirrer plate.

TEV cleavage

Pooled IMAC elution fractions were cleaved overnight at 4°C with an approximate equimolar concentration of His_6-TEV protease (S219V) with agitation on a rotating wheel, rolling platform or magnetic stirrer.

Reverse IMAC

The TEV-cleaved protein solution was either centrifuged briefly or filtered through a 0.80 µm syringe filter to remove precipitates formed during the overnight cleavage. In later purification attempts, IMAC elution buffer was added to the protein sample to adjust the solution to 5 mM imidazole. Supernatants were run over a 5 mL HisTrap Excel column at 3 mL/min, collecting the flow through. The column was then washed with IMAC wash buffer (0 mM or 5 mM imidazole) at 5 mL/min to collect residual cleaved protein that was not specifically bound to the column.

Ion Exchange Chromatography (IEX)

Protein solutions from reverse IMAC were diluted 10-fold in IEX LSB to achieve a final concentration of 30 mM NaCl. Proteins were loaded onto a MonoQ 10/100 GL column at 6 mL/min via a sample line. Proteins were eluted from the column in a linear gradient from 30 mM NaCl to 400 mM NaCl over 7.4 CV into 1 mL fractions at 4 mL/min. Protein elutions were analysed by SDS PAGE prior to centrifugal concentration and subsequent purification by size-exclusion chromatography.

Size Exclusion Chromatography (SEC)

Elution fractions containing Cdc20 protein were pooled and concentrated in a Vivaspin 20, 10 K MWCO centrifugal filter (Sartorius Stedim) at 8000×g, 10°C in a fixed angle rotor. A Superdex 75 Increase 10/300 GL column was equilibrated with SEC buffer and proteins were injected onto the column using a 500 µL capillary loop. Samples were separated at 0.75 mL/min over 1 CV. If the total sample volume after concentration was > 0.5 mL, multiple runs were performed. Eluted proteins were analysed by SDS PAGE prior to concentration. Protein was aliquoted, flash frozen in liquid N_2 and were stored at -80°C until required.

4.2.4 Preparation of Cdc20 for SPR

His_6-TEV-Cdc20 (161-477)

Protein eluted from the IMAC column were pooled and concentrated. Injections of 500 µL were run on a Superdex 75 Increase 10/300 GL column at 0.75 mL/min in 25 mM HEPES,

150 mM NaCl, 1 mM $MgCl_2$, 0.5 mM TCEP, 5% (v/v) glycerol, pH 7.5. Elutions were pooled and concentrated to 10 μM in Vivaspin 20, 10 K MWCO centrifugal concentrator (Sartorius Stedim). The pooled and concentrated protein was analysed by SDS PAGE. Cdc20 was then aliquoted and flash frozen in liquid N_2 before storing at -80°C, until required.

Biotinylation of Cdc20 (161-477)

The protocol for the biotinylation of protein was adapted from [219]. Cdc20(161-477) was thawed from -80°C and centrifuged at 20,000×g, 4°C for 10 minutes to remove any precipitates. Cdc20 (40 μL, 48 μM) was diluted in biotinylation buffer (10 mM Tris-HCl, 150 mM NaCl, 1 mM $MgCl_2$, 5% (v/v) glycerol, 1 mM TCEP, pH 7.3) to a final volume of 500 μL. A 1 mg aliquot of Sulfo-NHS-LC-LC-Biotin (Thermo Fisher Scientific, A35358) was dissolved to a 1 mM stock solution in biotinylation buffer. The Sulfo-NHS-LC-LC-Biotin solution was added to the protein to achieve a final molar ratio of 0.9:1 of biotin reagent to Cdc20. The mixture was vortexed briefly and allowed to react on ice for 3 hours. The sample was then purified on a Superdex 75 Increase 10/300 GL column in biotinylation buffer at 0.75 mL/min to remove the unreacted Sulfo-NHS-LC-LC-Biotin. Elution fractions were pooled, aliquoted and flash frozen in liquid N_2 before storing at -80°C.

4.3 Results and Discussion

4.3.1 Test expressions

The over-expression of Cdc20 in each baculovirus infection was assessed using two metrics. The cell count information provided qualitative information on successful viral infection in conjunction with protein visualisation on a SDS PAGE gel. Two constructs lengths of the $Cdc20^{WD40}$ domain were explored. The first was the Tian et al. published construct comprising residues 161-477 [193]. The second was a four amino acid truncation from the N-terminus of this construct, i.e. 165-477, which corresponds to the residues observed and modelled in the 4GGC PDB structure. Marginal Sf21 cell growth arrest was observed in the viruses containing MBP-tags. In addition, these constructs showed similar viabilities compared to the non-infected control, indicating poor infection. In contrast, cultures infected with Cdc20 viruses without the additional MBP-tags showed cell growth arrest, a drop in cell viability and cell swelling, indicating good viral infection and expression as was observed with the positive control sample, YFP (Table 4.3). In addition, samples from small-scale purifications using a PhyNexus instrument were analysed by SDS PAGE. As can be seen in Figure 4.3, SDS PAGE gel analysis provided a more direct measure of protein over expression. Viruses His_6-TEV-Cdc20 (161-477) and His_6-TEV-Cdc20 (165-477), lanes 3 and 5 respectively, appear to have an additional protein expressed around 40 kDa, compared to the non-infected control sample. Proteins at this molecular weight are in agreement with the expected molecular weight for the constructs (Table 4.1).

Table 4.3: Cell count information prior to harvesting and PhyNexus runs for each virus construct. YFP is used as a postive control virus (provided by AstraZeneca). Non-infected control gives baseline values for average cell diameter, % viability and indication on cell growth arrest in positive samples.

Sample	Final cell count	% viability	avg. diameter (μm)
His_6-MBP-TEV-Cdc20 (161-477)	60.85×10^5	92.5	21.83
His_6-TEV-Cdc20 (161-477)	34.21×10^5	87.4	22.74
His_6-MBP-TEV-Cdc20 (165-477)	91.62×10^5	92.5	19.74
His_6-TEV-Cdc20 (165-477)	37.77×10^5	87.0	24.01
YFP - positive control	29.85×10^5	70.5	22.76
Non-infected - negative control	102.53×10^5	94.2	19.61

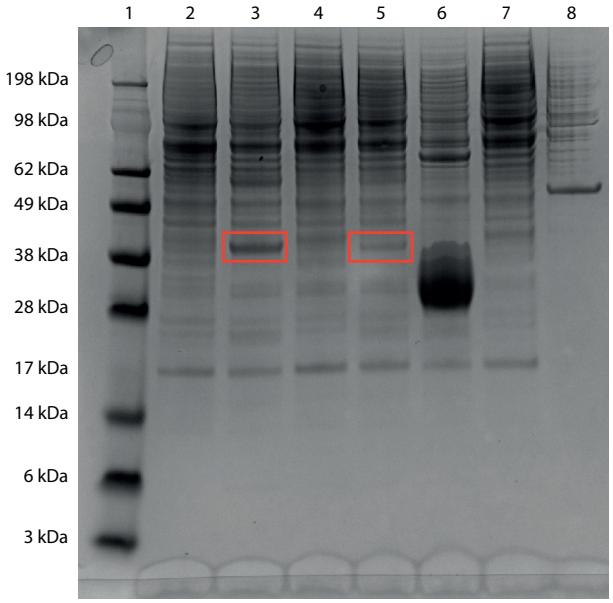


Figure 4.3: SDS PAGE gel of elutions fractions from Phynexus runs with Ni-NTA tips. Lane 1: SeeBlue Plus2 marker, lane 2: His_6-MBP-TEV-Cdc20 (161-477), lane 3: His_6-TEV-Cdc20 (161-477), lane 4: His_6-MBP-TEV-Cdc20 (165-477), lane 5: His_6-TEV-Cdc20 (165-477), lane 6: YFP, positive control, lane 7: Non-infected control, lane 8: BSA standard. Protein band highlighted in the red boxes corresponds to the expected molecular weight of the Cdc20 construct.

Due to unforeseen circumstances, the Cambridge insect cell culture facilities of AstraZeneca were moved to Macclesfield during the final year of the PhD. This further complicated the production of new insect cell pellets, since it was not possible to obtain the AstraZeneca Sf21 cell line for work outside of AstraZeneca at a time when the expression protocols had already been optimised for this particular cell line. As a consequence insect cell culture had to be performed within the Department of Pharmacology using other insect cell lines. Test expressions were performed in Sf9 and Hi5 cell lines, using the His_6-TEV-Cdc20 (161-477), due to its higher expression in the Sf21 cell line (Figure 4.3), compared to the His_6-TEV-Cdc20 (165-477) virus. Cell growth in the Hi5 cells arrested, but this did not appear to occur in the Sf9 cells. In addition, the Sf9 cell viability did not drop to the same extent as in the Hi5 cells. However, the average diameter of the Sf9 cells did increase significantly, whereas it appeared to decrease in the Hi5 cells (Table 4.4). The drop in average diameter in the Hi5 cells may be artefactual due to a change in cellular morphology from circular to an elongated shape, which was often counted as multiple smaller cells by the LUNA-II cell counter. Despite the initial appearance from the cell count data, Figure 4.4 shows that the Sf9 cells are expressing the construct at a higher level compared to the Hi5 cells. All large-scale expression cultures were then performed with Sf9 cells.

Table 4.4: Cell count information prior to harvesting and purification runs for Cdc20. Non-infected control gives a baseline of average cell diameter, % viability and indication on cell growth arrest in positive samples.

Sample	Final cell count	% viability	avg. diameter (μm)
Sf9 - infected	3.89×10^6	93.0	15.1
Sf9 - non-infected	4.68×10^6	98.1	11.3
Hi5 - infected	1.71×10^6	86.6	15.3
Hi5 - non-infected	3.35×10^6	94.0	16.9

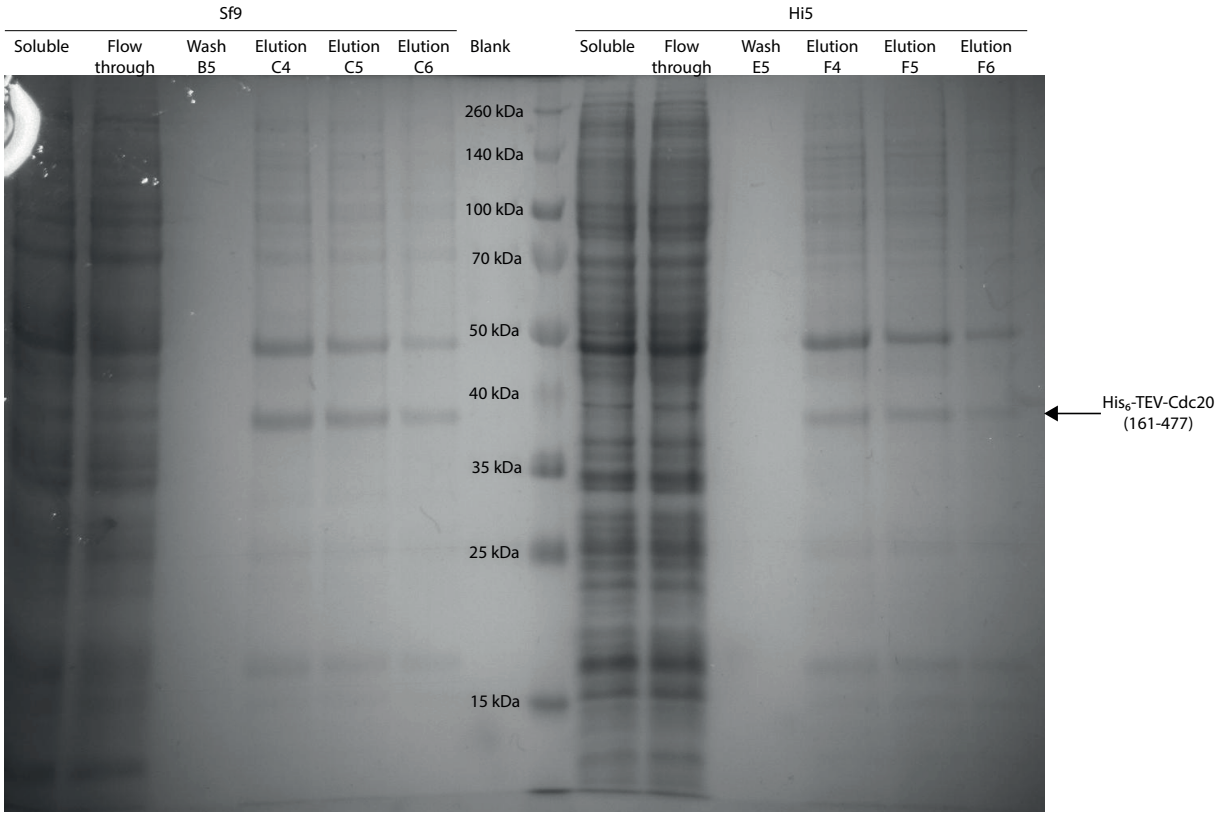


Figure 4.4: SDS PAGE gel of purification fractions from 1 mL HisTrap Excel column. Protein band highlighted at arrow indicator, correspond to proteins of the expected molecular weight for the Cdc20 construct.

4.3.2 Purification of Cdc20^{WD40} protein

Every purification attempt differed in its protocol to some extent, where gradual improvements were attempted at each iterative chromatography stage. The result was that no two purifications were alike. Nevertheless, at each stage, progress was made and pure Cdc20 protein was obtained, but yield was often diminished. In each chromatography step, SDS PAGE gels are shown, highlighting where chromatography was not optimal and how it was improved accordingly.

Cell lysis

It was assumed that optimal lysis could be performed using the C5-Emulsiflex as was normally used for the lysis of *E. coli* cells. However, unlike *E. coli*, insect cells lysate is very viscous causing air bubbles to be trapped during lysis that could not be removed using high-speed centrifugation. The air bubbles and viscosity can create downstream problems during liquid chromatography, as both air and high pressure due to very viscous liquids could damage chromatography columns. However, unlike *E. coli*, insect cells membranes are significantly weaker and can be lysed by a simple freeze-thaw cycle. The supernatant from freeze-thaw lysis was considerably less viscous than when lysed using the C5-Emulsiflex, and did not contain air trapped within it. Initially, freeze-thaw lysis was performed using re-sealable bags on dry ice for total volumes of approximately 100 mL. However, in scaled-up protein purifications (cell pellets from 3.5 L of expression culture) in which lysate volumes were around 300 mL, this process became less feasible. Instead, resuspended lysates were aliquoted into 15 mL tubes and were freeze-thaw lysed in liquid N_2, due to ease of handling.

IMAC

Standard buffer compositions (Table 4.2, first generation) and protocols were used in early IMAC purifications. Proteins bound to the 5 mL HisTrap Excel column were washed with 20 CV of IMAC wash buffer and eluted from the column. In early purification attempts, the final wash fraction (wash 7 in Figure 4.5a) clearly indicates that the His_6-TEV-Cdc20 (161-477) protein was eluting prematurely from the column. As a result, the yield from this chromatography step was particularly poor. In subsequent purifications, the concentration of imidazole in the IMAC wash buffer was decreased to 15 mM, however premature elution from the IMAC column was still observed. To reduce this loss of protein, the concentration of imidazole was further reduced to 10 mM and the buffer pH was increased to 8.5 yielding the optimal second-generation buffers listed in Table 4.2. These alterations resulted in a better overall yield from this chromatography step, but also came at a cost of eluent purity, as can be seen in Figures 4.5b and 4.5c. Retrospectively, it must be noted that the imidazole supplier was changed mid-way through these purifications. The latter batch of imidazole was of lower purity than the previous batch, which could be observed by its higher A280 signal and more distinct off-yellow colour compared to the previous batch. Furthermore, other individuals in our laboratory had also observed their His_6-tagged constructs washing off at lower concentrations of imidazole than expected. Thus, removal of protein contaminants was less efficient compared to purification using higher purity imidazole. Unfortunately due to the limiting nature of reagents, it was not possible to systematically determine what was the true cause of premature elution from

the IMAC column.

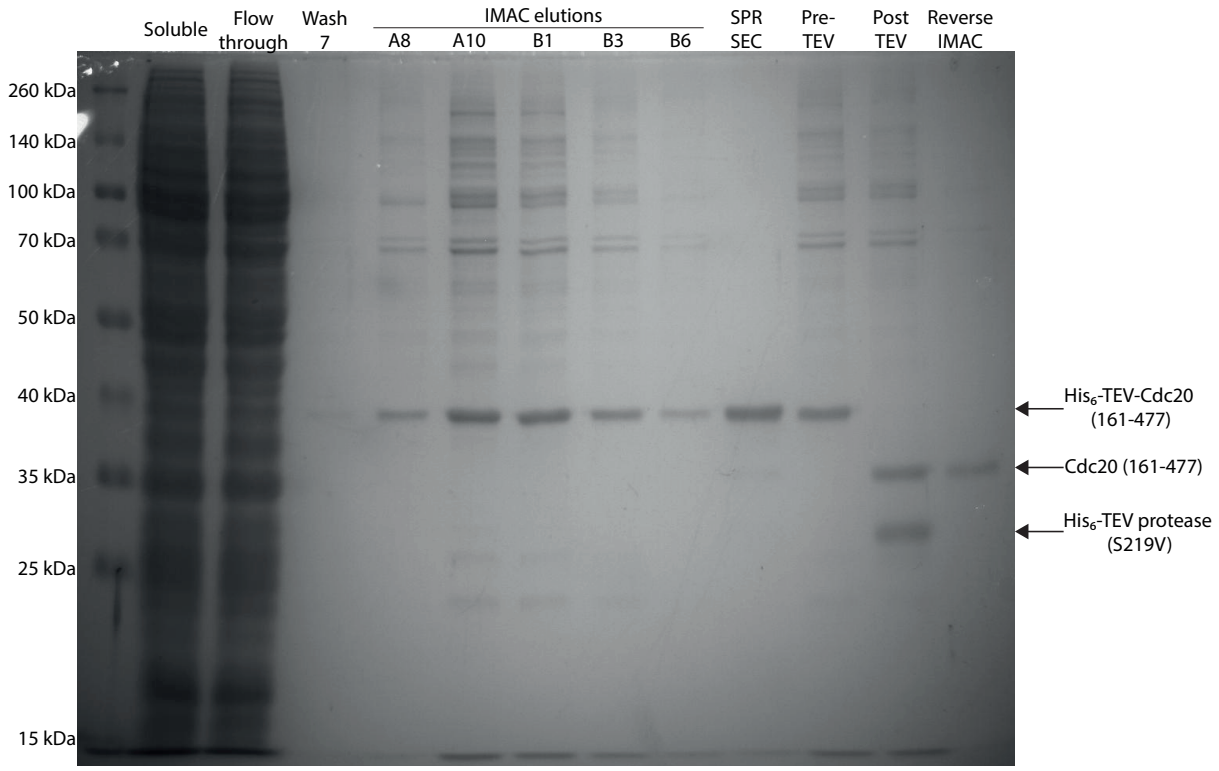
Protein precipitation directly after IMAC elution into a 96 deep-well block was the major source of protein loss. Precipitation would occur in a concentration-dependant manner as could be observed when comparing the solution in the 96 deep-well block to the ÄKTA chromatograms. Since samples were always taken for SDS PAGE gel analysis after centrifugation, it was not possible to know how much Cdc20 was lost. However, given the low yield often obtained, every step to improve the yield had to be explored. In later IMAC purifications, alternative methods were employed to reduce precipitation. Since protein precipitation appeared concentration-dependent, immediate dilution of the IMAC eluent should be beneficial. The first attempt was to elute into 50 mL tubes pre-filled with volumes of IMAC wash buffer (without imidazole). While the principle did work and significantly less precipitation was observed, it was realised that physical mixing was required during elution. Since this was not possible, using the fraction collector on the ÄKTA, an alternative had to be developed. Elution via the outlet valve solved this issue, where it was possible to directly connect the line from the outlet valve to a duran bottle containing IMAC wash buffer (without imidazole) placed on a magnetic stirrer plate. With gentle stirring during IMAC elution, no precipitation was observed.

TEV cleavage

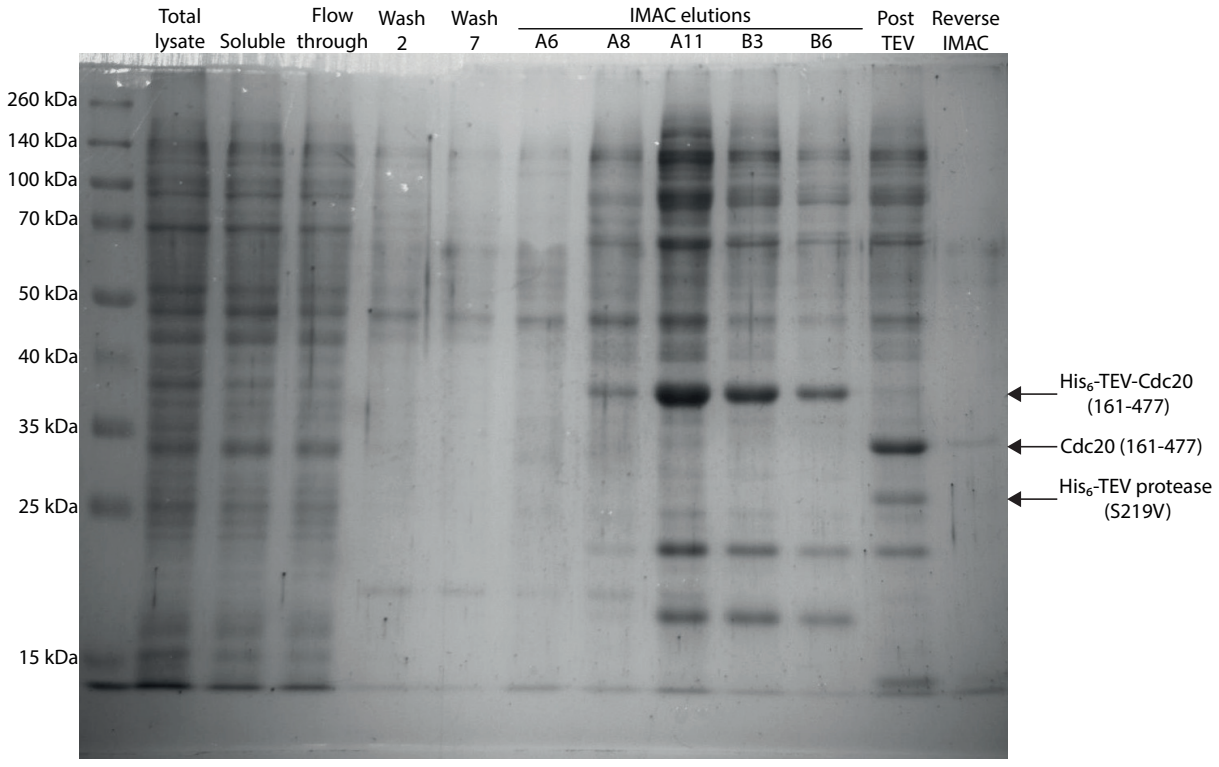
An approximate equimolar amount of His_6-TEV protease (S219V) was used for the overnight cleavage of the His_6-TEV-Cdc20 (161-477) protein. This high concentration of TEV protease was used for two main reasons. The specific activity of the enzyme after purification (Section 2.3.3) was never calculated. In addition, the following chromatography step would remove all His_6-tag containing proteins on the IMAC column, thereby separating the excess TEV protein from the desired product. Since incomplete cleavage of the Cdc20 construct would result in a significant loss of yield, an excess of TEV protease was always added to ensure complete cleavage.

Reverse IMAC

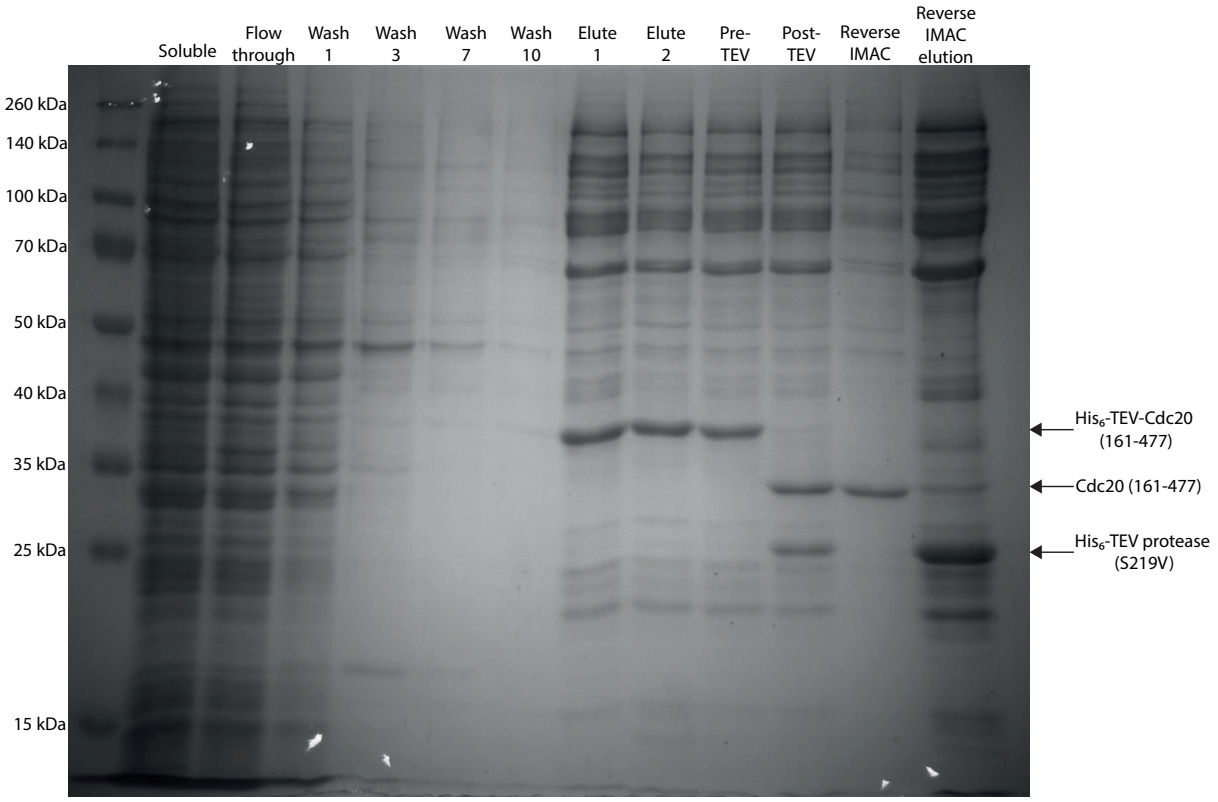
IMAC purification from insect cells typically contains several contaminating proteins at high molecular weights that bound non-specifically to the HisTrap column. These contaminants should be able to re-bind the IMAC column in the same manner, and this could serve as an effective purification strategy. In all reverse IMAC purifications, re-binding of contaminants was observed to varying degrees (Figure 4.5). In some instances, Cdc20 without the His_6-tag was also observed to re-bind the IMAC column, likely due to its low solubility and 'sticky' nature. In cases where this occured, it was possible to re-elute


Soluble
Flow through
Wash 7
IMAC elutions
A8
A10
B1
B3
B6
SPR SEC
Pre-TEV
Post TEV
Reverse IMAC
260 kDa
140 kDa
100 kDa
70 kDa
50 kDa
40 kDa
35 kDa
25 kDa
15 kDa
His6-TEV-Cdc20 (161-477)
Cdc20 (161-477)
His6-TEV protease (S219V)

(a)


Total lysate
Soluble
Flow through
Wash 2
Wash 7
IMAC elutions
A6
A8
A11
B3
B6
Post TEV
Reverse IMAC
260 kDa
140 kDa
100 kDa
70 kDa
50 kDa
40 kDa
35 kDa
25 kDa
15 kDa
His6-TEV-Cdc20 (161-477)
Cdc20 (161-477)
His6-TEV protease (S219V)

(b)



(c)

Figure 4.5: SDS PAGE gels of IMAC, TEV cleavage and reverse-IMAC purification steps. (a) IMAC column was washed with using first-generation buffer (Table 4.2). (b) IMAC column washed with 20 CV, second-generation IMAC buffer. (c) IMAC column was washed with 30 CV second-generation buffer.

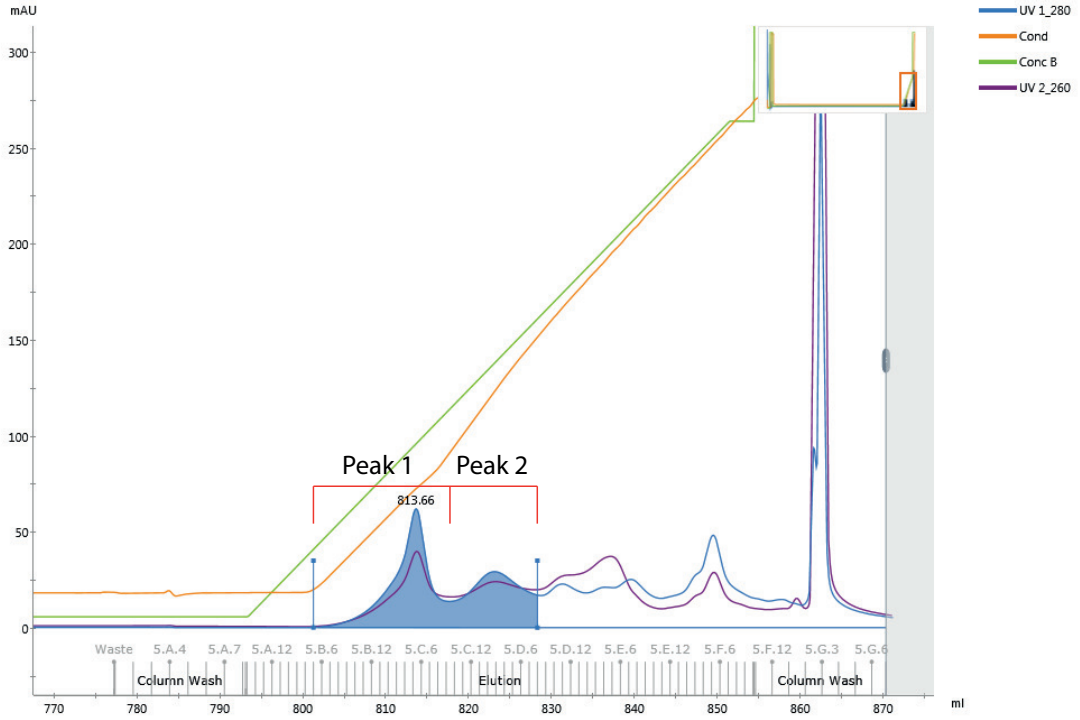
Cdc20 using IMAC elution buffer, however the eluent would inevitably be more contaminated and due to the high amount of impurities, it was not possible to further purify this protein. In later purifications, non-specific binding of Cdc20 was prevented by the addition of IMAC elution buffer into the TEV cleaved sample to a final concentration of 5 mM imidazole. This addition drastically improved the recovery of Cdc20 in the reverse IMAC flow through fraction, but also came at a cost of purity since varying amounts of the high molecular weight contaminants were also prevented from non-specific interaction with the column (Figure 4.5).

IEX

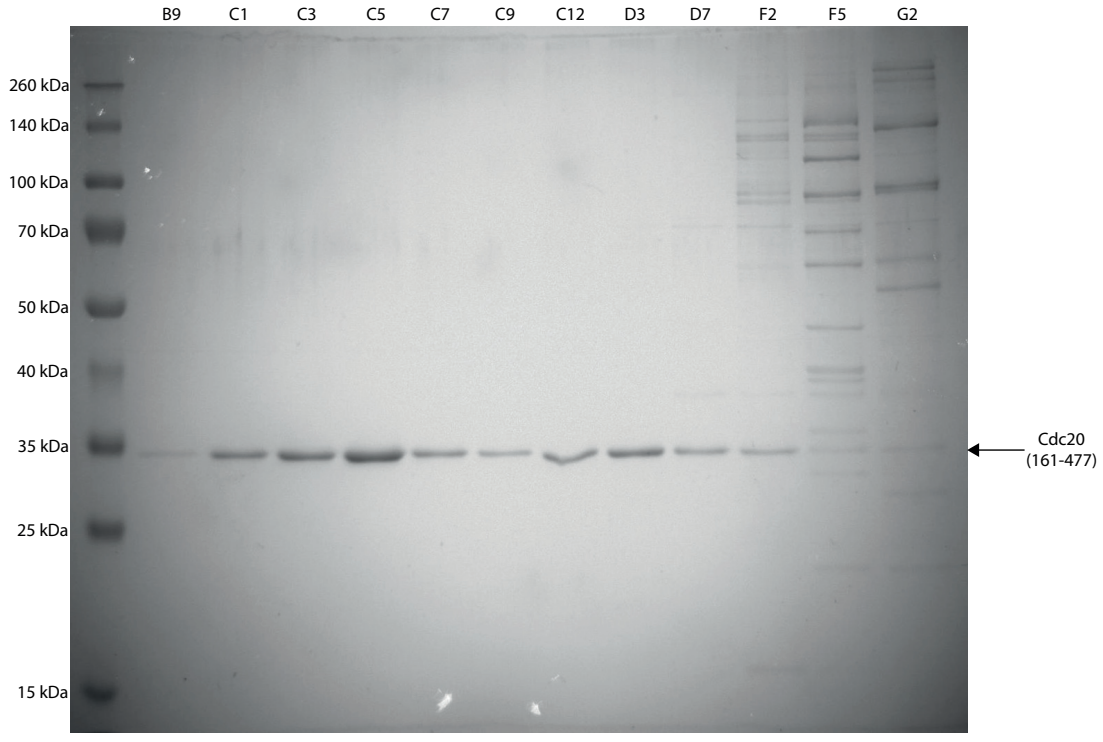
Ion exchange chromatography is often used to separate molecules of similar molecular weights, with varying charge and is often employed as an intermediate purification step between affinity chromatography and SEC. The theoretical isoelectric point (pI) of Cdc20 (161-477) was calculated at 6.58 using Protparam [202]. In all IEX buffer systems (Table 4.2), Cdc20 should be net negatively charged and able to bind anion exchange resins, such as the MonoQ (GE Healthcare). Proteins were diluted 10-fold in MonoQ LSB to achieve a concentration of 30 mM NaCl to enable proteins to bind the MonoQ 10/100 GL column (GE Healthcare). Due to the typically large elution volumes from the IMAC column, the resulting volume to be loaded onto the anion exchange columns was typically between 0.5 - 1.5 L. As a consequence proteins were loaded at the maximal flow rates allowed for the column (dependant on the pre-column pressure limits) to reduce loading times. Nevertheless, sample loading would often span hours.

All proteins loaded onto the MonoQ 10/100 GL column would elute before the gradient reached 400 mM NaCl. Most contaminant proteins eluted from the column towards the end of the gradient, while Cdc20 would elute earlier, resulting in highly pure Cdc20 from the MonoQ column as monitored by SDS PAGE (Figure 4.6b). During IEX chromatography proteins typically elute in a sharp peak. However, in the case of Cdc20, elution was always observed to be biphasic with broad peaks occurring between 50 mM and 250 mM NaCl, of which Peak 1 was right-skewed and peak 2 was symmetric (Figure 4.6a). This was observed regardless of the buffer pH used, i.e pH 7.5 or pH 8.5, and the presence or absence of 5% glycerol. Despite the elution profile being non-standard for Cdc20, the technique was still used as the sample purity would be increased. Due to the typically low yields of protein, the two separate peaks were never investigated independently to look at any functional differences.

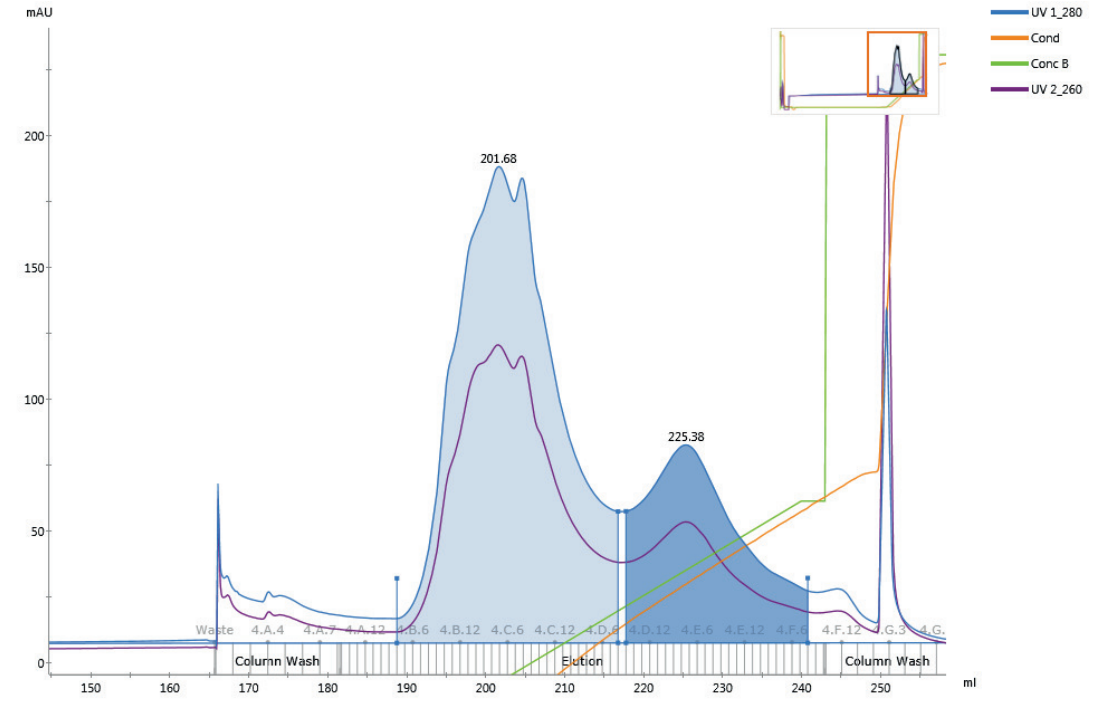
In reverse IMAC purifications where Cdc20 was observed to non-specifically bind the IMAC column, purification of this sample was attempted by IEX. However, due to the increased number of contaminant proteins, only poor separation was observed (Figures 4.6c and 4.6d). This highlighted the necessity to first obtain the highest possible purity of protein using IMAC.


mAU
300
250
200
150
100
50
0
UV 1_280
Cond
Conc B
UV 2_260
Peak 1
Peak 2
813.66
Waste
5.A.4
5.A.7
5.A.12
5.B.6
5.B.12
5.C.6
5.C.12
5.D.6
5.D.12
5.E.6
5.E.12
5.F.6
5.F.12
5.G.3
5.G.6
Column Wash
Elution
Column Wash
770
780
790
800
810
820
830
840
850
860
870
ml

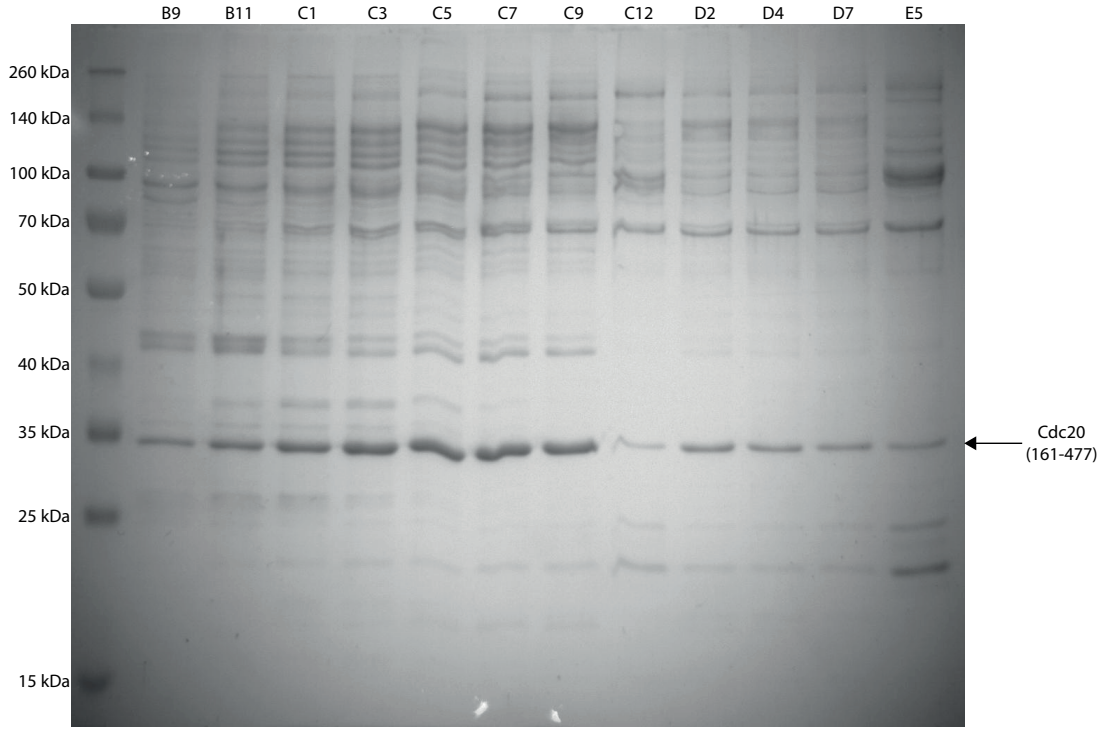
(a)


B9
C1
C3
C5
C7
C9
C12
D3
D7
F2
F5
G2
260 kDa
140 kDa
100 kDa
70 kDa
50 kDa
40 kDa
35 kDa
25 kDa
15 kDa
Cdc20
(161-477)

(b)



(c)



(d)

Figure 4.6: ÄKTA chromatograms and their associated SDS PAGE gels of anion exchange chromatography purifications. (a, b) Sample from reverse IMAC flow through. (c, d) Sample from reverse IMAC re-elutions.

SEC

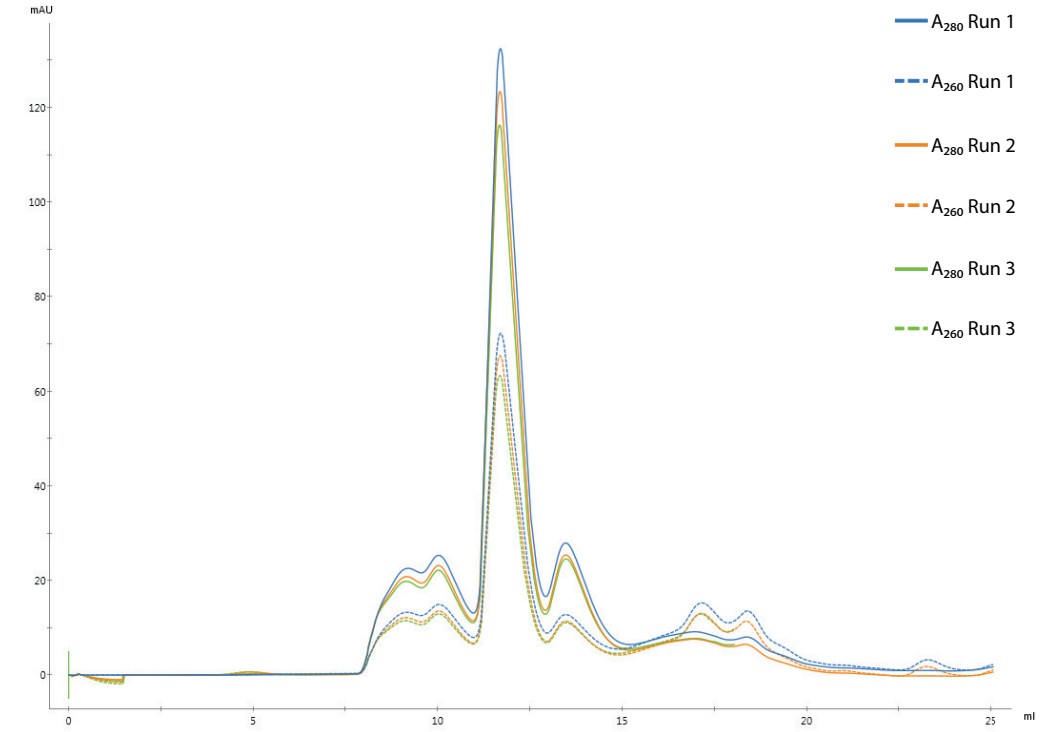
Due to the broad elution profiles from anion exchange chromatography, elution fractions had to be concentrated prior to SEC. Centrifugal concentration would often take several hours to achieve volumes ranging from 1 mL to 3 mL of protein. Slow concentration times are indicative of protein precipitation occurring on the membrane. In order to reduce the build up of concentration gradients within the concentrator, protein solutions were resuspended at regular intervals. Precipitates in solution was rarely observed, but it is possible that they would remain bound to the concentrator membrane.

A Superdex 75 Increase 10/300 GL column (GE Healthcare) was used size exclusion chromatography, due to the superior resolution and higher mechanical stability of the Increase matrix compared to the standard Superdex matrix. This allowed the application of flow rates of 0.75 mL/min, without losing significant resolution resulting in run times of less than 30 minutes. Even with complex sample inputs, good separation could be observed (Figure 4.7b). The column performed very well over multiple, consecutive and non-consecutive runs, as is highlighted in the very reliable retention times for the elution of Cdc20 (Figure 4.7a). This reliability aided the precise fraction collection and monitoring the standard of protein purity between different preparations.

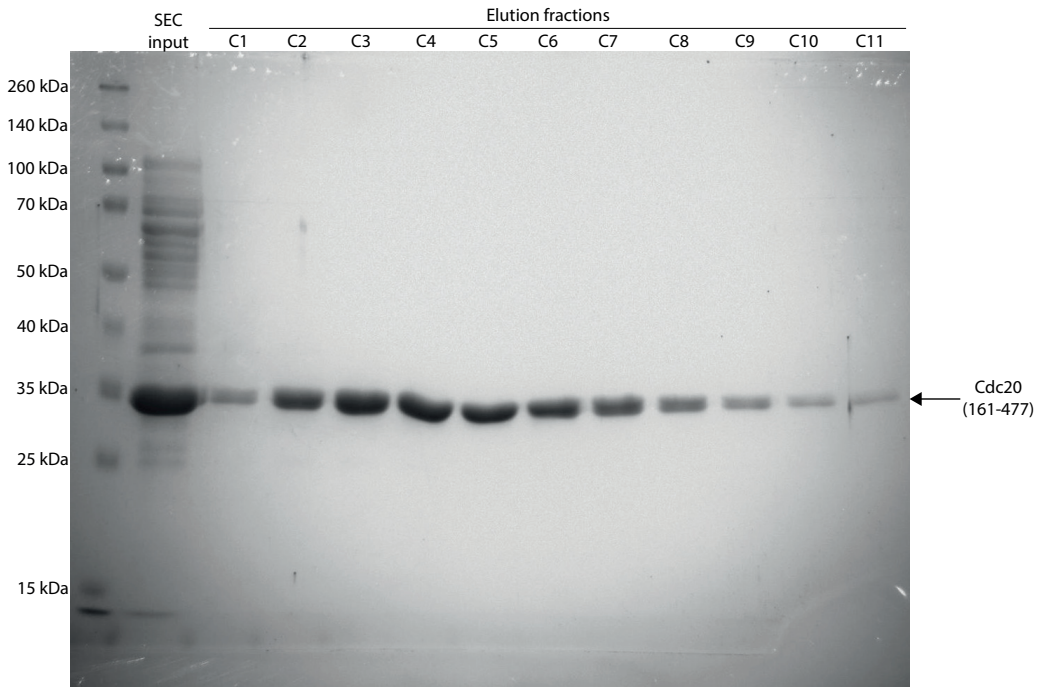
Purified Cdc20 was to be used in three main downstream experiments. Both methods for biophysical characterisation of peptide ligands (DSF and SPR) rarely required any concentration after elution from the SEC column. However, for X-ray crystallography, centrifugal concentration was necessary. In all instances, we observed similarly long concentration times as described earlier, and protein from multiple preparations had to be pooled in order to generate enough sample to prepare a single 96-well plate for a crystallization trail.

4.3.3 Preparation of Cdc20 for SPR

Two ligand immobilisation techniques were explored for SPR. In both methods, an affinity capture approach was preferred over standard amine coupling that requires preconcentration using low pH buffer, in which many proteins are unstable. The first method utilised the affinity of His_6-tagged Cdc20 construct to a Ni-NTA biosensor chip. The second method utilised biotinylated-Cdc20 binding to a streptavidin-coated biosensor chip. An *in vitro* biotinylated construct was considered utilising the Avi-tag and BirA biotin ligase for biotinylation site specificity. However, given the need to clone and generate new baculovirus it was considered unnecessary for the small scale of protein that would inevitably be required for immobilisation onto the biosensor.



(a)



(b)

Figure 4.7: (a) SDS PAGE gel of Cdc20 purification using a Superdex 75 Increase 10/300 GL column. Input sample run on gel shows contaminant proteins that are separated from the Cdc20 in the elution fractions. (b) Overlay of multiple ÄKTA chromatograms monitoring the A_{280} (solid line) and A_{260} (dashed line) of multiple sample injections onto the column. Blue = run 1, orange = run 2, green = run 3.

His_6-TEV-Cdc20 (161-477)

Despite the normal Cdc20 purification requiring IEX chromatography, it was possible to separate contaminant proteins present in a higher purity IMAC elution sample using the Superdex 75 Increase 10/300 GL column equilibrated in HEPES buffer since the protein would be amine coupled to the Ni-NTA biosensor after affinity capture. The purity of this sample was confirmed by SDS PAGE of the pooled and concentrated protein and was estimated to be >95% pure (Figure 4.5a, lane 'SPR SEC').

Biotinylation of Cdc20 (161-477)

The EZ-link^{TM} Sulfo-NHS-LC-LC-Biotin (Thermo Fisher Scientific) was used for chemical biotinylation. This particular reagent was used due to the length of the linker provided between the biotin group and Cdc20. Despite the commonly assumed principle that NHS/EDC reactions should not be performed in buffers containing primary amines, Papalia and Myszka present data that prove otherwise and show the reaction to work in Tris-HCl buffer at a pH below the pK_a of Tris. Due to the possibility of losing protein during buffer exchange, Cdc20 was instead diluted in a buffer based on the results from Papalia and Myszka [219]. The conjugation stoichiometry was determined as 1:1 of biotin:Cdc20, and the yield was estimated to be $\approx$ 10% by ESI-MS on a Synapt G2-Si (Waters), run by Dr. Christopher Stubbs at AstraZeneca. The predominant mass observed corresponded to that of unconjugated Cdc20 (34847.1 $\pm$ 0.1 Da, Figure 4.8b), which agrees with the predicted molecular weight (Table 4.1). A second mass corresponding to Cdc20 coupled to a single Sulfo-NHS-LC-LC-Biotin molecule (35299.2 $\pm$ 0.4 Da) was detected. This mass shift of 452.1 Da is in agreement with the expected addition of 452.6 Da [220].

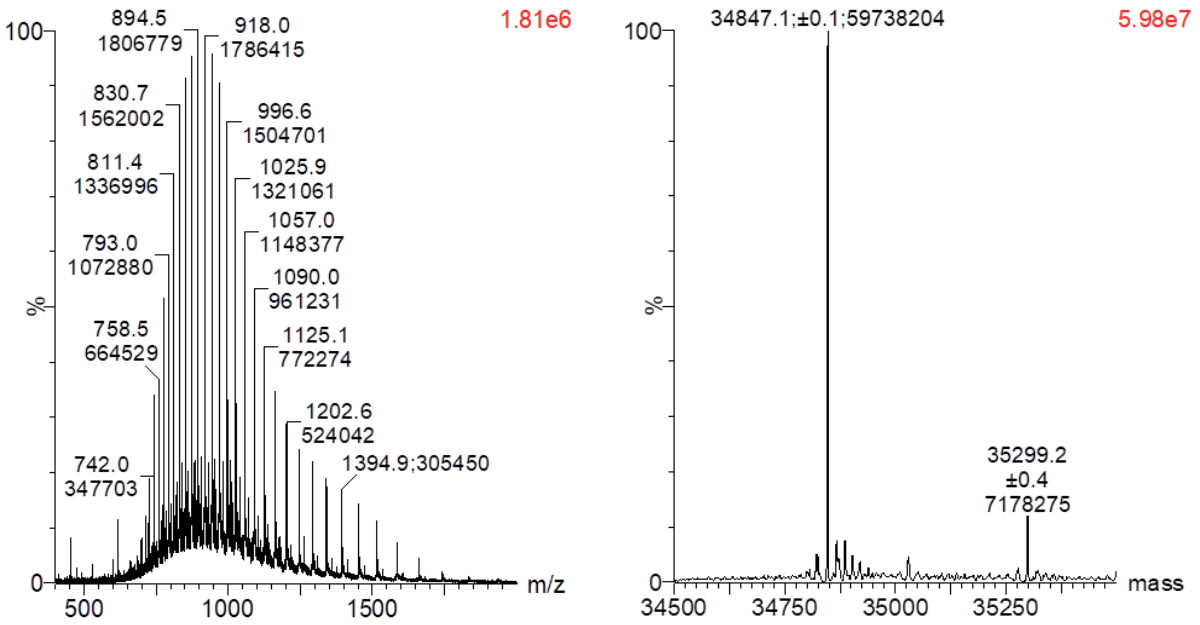


Figure 4.8: (a) m/z spectra of 25 nM Cdc20/ minimally biotinylated Cdc20. (b) Calculation of derived total masses using the MassLynx, maximum entropy 1 (MaxENT1) calculation from the m/z spectra.

4.4 Conclusions

This chapter discusses the production of Cdc20, which was used downstream to characterise synthesised peptide ligands. First, a small series of Cdc20 constructs, based on previously reported data described in the literature, was generated. For each construct, recombinant baculovirus was produced by transfection in Sf21 insect cells. Out of the four baculoviruses produced, the two viruses without the MBP-tags expressed the desired proteins at levels detectable by SDS PAGE. Cdc20 (161-477), the literature precedent construct published by Xuelian Luo's lab [193, 195], was observed at a higher yield per mL of Sf21 culture compared to the shorter construct.

One of the most challenging aspects of this project has been to produce suitable quantities of Cdc20 for biophysical and structural characterisation. In order to obtain pure protein in sufficient quantities, purification methods were iteratively adapted between each purification attempt to improve the yield. Generating a good protocol for IMAC was particularly important in obtaining highly pure protein, but precipitation at this step was a significant issue and was likely the cause for reduction in the overall yield. IEX chromatography was also a necessary step in the majority of purifications, however an op-

timal elution profile in which Cdc20 would elute in a sharp peak from the MonoQ column was never obtained. SEC using a Superdex 75 Increase 10/300 GL column was employed as a final polishing step that resulted in highly pure protein. Solubility of Cdc20 was always an issue, particularly when attempting to concentrate the protein using centrifugal concentrators in preparation for SEC or in achieving concentrations suitable for setting up protein crystallisation plates.

Biotinylation of Cdc20 for use in SPR experiments was also explored. The amine coupling reaction was performed in a buffer containing a primary amine (Tris), which is contrary to what is standard in the field. Nevertheless, single-site biotinylation of Cdc20 was observed by ESI-MS and the final yield, although only 10%, was more than sufficient for future SPR experiments.

Chapter 5

Design and synthesis of APC/C^{Cdc20} peptide inhibitors

5.1 Introduction

This chapter will focus on the process of developing peptides against the APC/C and its co-activator Cdc20. Here, the aim is give an overview of the peptides before describing their binding activities in the following chapter. The design of peptides was based on knowledge from consensus degron sequences in addition to structural data available from EM and crystal structures of Cdc20 and Cdh1 bound to peptides derived from various APC/C substrates. Only limited structural information is available for degron motifs bound to Cdc20. However, Cdc20 and Cdh1 share a high degree of structural homology, particularly around the WD40 domain containing the degron-binding sites. Therefore, data from both Cdc20 and Cdh1 were used in the design process. Initially, a small subset of peptide inhibitors were designed and synthesised targeting the binding sites for D-box, KEN-box and ABBA motifs (Figure 5.1). Additionally, peptides were designed to inhibit interactions between Cdc20 and the APC/C by mimicking the N-terminal C-box motif of Cdc20.

5.2 Methods

5.2.1 Cryo-EM and X-ray crystal structure analysis

Models of Cryo-EM or X-ray crystal structures were analysed using PyMOL [204]. Analysis of electron density maps (2Fo-Fc) was performed in some instances and map coefficients file (.mtz) were downloaded from RCSB PDB [221, 222]. Coordinate (.pdb) and map coefficients files (.mtz) were loaded into Coot [213] to assess the 2Fo-Fc data in local areas,

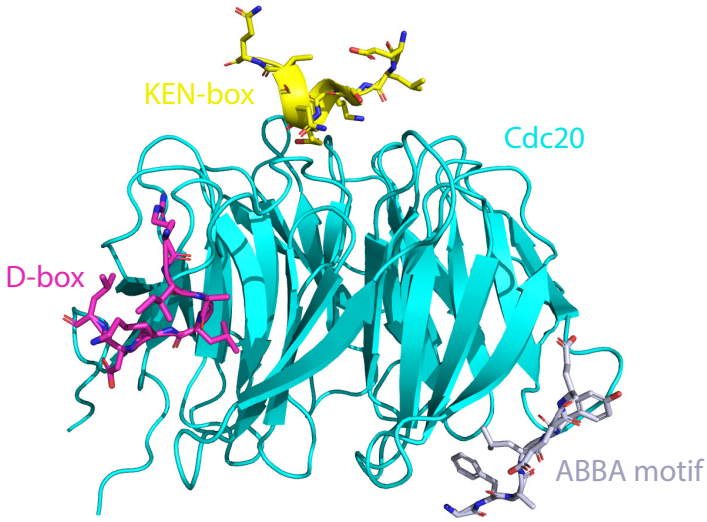


Figure 5.1: Overlay of the D-box (magenta, PDB:4BH6 [165], the KEN-box (yellow, PDB:4GGD [193]) and the ABBA motif (grey, PDB:4BH6 [165]) with Cdc20 (cyan, PDB:4GGD [193].

contoured at 1.5σ.

5.2.2 Solid-phase peptide synthesis (SPPS)

All peptides were synthesised using a vacuum manifold setup to aid the parallel synthesis of multiple peptides simultaneously (Figure 5.2). Peptides were synthesised on Ramage-ChemMatrix® resin (Sigma Aldrich, 727792), using Fmoc-protected amino acids (Sigma Aldrich). Fmoc-L-cBuAla-OH (CAS no: 478183-62-9), Fmoc-L-NptGly-OH (CAS no: 139551-74-9) and (s)-Fmoc-2-amino-4,4,4-trifluoro-butryic acid (CAS no: 181128-48-3) were purchased from Iris Biotech GmbH. Hexafluorophosphonate azabenzotriazole tetramethyl uronium (HATU) and 1-Hydroxy-7-azabenzotriazole (HOAt) were purchased from ApexBio and Sigma Aldrich, respectively. Peptides were synthesised using two protocols, of which the second protocol was optimised to increase the throughput of coupling reactions and save reagents. The second method did not appear to compromise the yield or purity of the crude sample.

Method 1

Synthesis was performed on a 0.1 mmol scale. Ramage-ChemMatrix® resin (204 mg, loading capacity of 0.49 mmol/g) was weighed into a Bio-rad Econo-Pac® polypropylene chromatography column with a nylon frit at the base. All swelling, coupling, de-protection and washing steps of resins were performed with gentle bubbling using compressed air or N_2. Prior to coupling, resins were swelled for at least 30 minutes in DMF. The Fmoc-protecting group from the previously coupled amino acid was de-protected in 20 % (v/v)

Figure 5.2: SPPS using polypropylene columns on a vacuum manifold set up. Resins are bubbled gently to ensure mixing from an inlet line of compressed air or N_2, through a series of 3-way luer-lock valves connected in parallel. Valve positions allow bubbling, drainage or capping. Low vacuum pressure aids the drainage of solutions in the column.

piperidine in Dimethylformamide (DMF) for 15 minutes. Following de-protection, resins were washed five times in DMF to remove residual piperidine. Meanwhile, Fmoc-L-amino acids (5 molar equivalents (eq.)), HATU (5 eq.) and HOAt (5 eq.) were dissolved in 2 mL of *N*-Methyl-2-pyrrolidone (NMP). *N*,*N*-Diisopropylethylamine (DIEA, 8.5 eq.) was added to activate the resulting mixture for 7 minutes. Activated Fmoc-L-amino acids were coupled for 1 hour to the resin were washed three times with DMF to remove uncoupled amino acids. This process was repeated for each amino acid to be coupled to the peptide. Finally, peptides were N-terminally acetylated in a 10 mL cocktail of acetic anhydride, DMF and DIEA (2:2:1), respectively for 1 hour.

Method 2

Synthesis was performed on a 0.1 mmol scale using Ramage-ChemMatrix® resin with the following alterations to the protocol. Fmoc-L-amino acids (2 eq.), HATU (2 eq.), HOAt (2 eq.) were dissolved in ≈ 2 mL of NMP. DIEA (3.4 eq.) was added to activate the resulting mixture for 1 minute. Activated Fmoc-L-amino acids were coupled for 10 minutes. Coupling reactions of Fmoc-L-Arginine were performed with twice with 5 eq. for 30 minutes each. In addition, the N-terminal acetylation reaction was reduced to 10 minutes.

5.2.3 Ring-closing olefin metathesis

Resins were washed 4 times in DMF, 3 times in MeOH and twice in Et_2O to clean and dry the resins prior to ring-closing olefin metathesis (RCM). All RCM reactions were performed in anhydrous dichloroethane (DCE). Resins were swelled in DCE while bubbling with N_2 gas before the reaction was intiated by the addition 3 mL of 10 mg/mL Grubb's 1st generation catalyst in DCE (Sigma Aldrich), bubbling in N_2 gas for 2 hours. The RCM reaction was repeated twice more with fresh reagent to increase the yield. The resins were then washed four times in 50% DMF/DMSO and were soaked overnight in the final wash solvent. An example of the RCM reaction is illustrated in Figure 5.3.

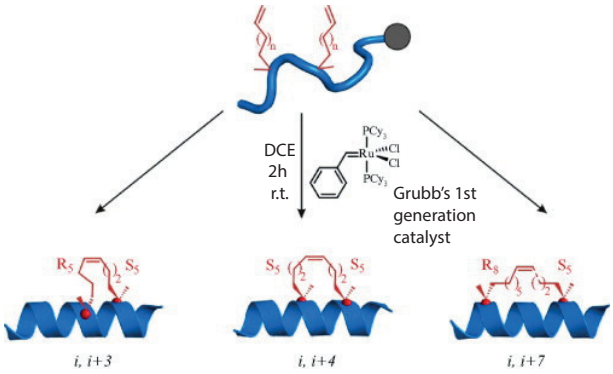


Figure 5.3: Illustration of the RCM reaction for i,i+3, i,i+4 and i,i+7 staples in α-helical peptides. Abbreviations: DCE - Dichloroethane, r.t. - room temperature, S_5 - *S*-2-(4'pentenyl) alanine, R_5 - *R*-2-(4'pentenyl) alanine and R_8 - *R*-2-(7'octenyl) alanine. Figure adapted from Verdine and Hilinski [223]

5.2.4 Peptide cleavage

Resins were washed 4 times in DMF, 3 times in MeOH and twice in Et_2O to clean and dry the resins appropriately and were then transferred to a 50 mL falcon tube. For each peptide, 8.6 mL of a cleavage cocktail containing 93% trifluoroacetic acid (TFA), 3.5% ddH_2O and 3.5% triisopropylsilane (TIPS) were added to the resin. Tubes were capped and incubated with vigorous shaking for 1 hour to cleave the peptide from the resin. Following this incubation, resins were separated from the cleaved peptide solution by filtration through the polypropylene chromatography column used in the synthesis into a fresh 50 mL falcon tube. Compressed air flow was then applied to the chromatography columns through an adapted lid to remove residual peptide solutions from the resins. Peptide solutions were triturated by the addition of 45 mL of Et_2O and the resulting peptide precipitate was isolated by centrifugation at 10,000 $\times$ g for 1 minute in a fixed angle rotor before drying under a flow of N_2. Peptides were then dissolved in a suitable

ratio and volume of ddH_2O/acetonitirile (MeCN) required to achieve full dissolution.

5.2.5 Analytical characterisation of peptides

Molecular weights of synthesised peptides were calculated using ChemDraw Ultra 12.0 software (CambridgeSoft, PerkinElmer). All peptide solutions were filtered through a 0.45 µm polytetrafluoroethylene (PTFE) syringe filter into a 250 µL polypropylene vial (Agilent) for Liquid chromatography mass spectrometry (LCMS) and High-performance liquid chromatography (HPLC) analyses.

Liquid chromatography-mass spectrometry (LCMS)

Chromatographs were recorded using a Waters ACQUITY H-Class UPLC with an ESCi Multi-Mode Ionisation Waters SQ Detector 2 spectrometer using MassLynx 4.1 software. Liquid chromatography runs separated on an ACQUITY UPLC® CSH C18 (2.1 mm × 50 mm, 1.7 µm, 130 Å) at 40°C, with a PDA eλ detector 220 – 800 nm, interval 1.2 nm. The following solvent and gradients were used for the runs. Solvent A: 2 mM NH_4OAc in 95% H_2O, 5% MeCN; solvent B: 100% MeCN; solvent C: 2% formic acid. The linear gradient was run from 5 - 95 % B with a constant of 5 % C over 1 minute, at a flow rate of 0.6 mL/min.

Analytical High-performance liquid chromatography (HPLC)

Analytical HPLC chromatographic separations were obtained on an Agilent 1260 Infinity system using a Supercosil ABZ+PLUS column (150 mm × 4.6 mm, 3 µm), eluting with a linear gradient system (solvent A: 0.1% (v/v) TFA in H_2O, solvent B: 0.05% (v/v) TFA in MeCN) over 15 min at a flow rate of 1 mL/min. Eluents were monitored by UV absorbance at 220 and 254 nm.

5.2.6 Semi-preparative HPLC

Semi-preparative HPLC was performed using an Agilent 1260 Infinity on a Supercosil ABZ+PLUS column (250 mm × 21.2 mm, 5 µm) eluting peptides with a linear gradient system (solvent A: 0.1% (v/v) TFA in H_2O, solvent B: 0.05% (v/v) TFA in MeCN) over 20 min at a flow rate of 20 mL/min. Eluents were monitored by UV absorbance at 220 and 254 nm.

5.2.7 Quantification of peptide mass

Purified peptide solutions from semi-preparative HPLC were lyophilised using a -100°C freeze dryer (LaboGene, CoolSafe) in 50 mL falcon tubes. Solutions were frozen in liquid

N_2 and placed under vacuum for solvent sublimation to occur. Lyophilised peptides were weighed using a 5 figure balance (0.01 mg resolution). Peptides were then dissolved to a desired concentration (e.g 5 mg/mL) in an appropriate solution of H_2O/MeCN to achieve complete dissolution. Peptide solutions were then aliquoted in equal volumes into 1.5 mL tubes and freeze-dried again to obtain multiple aliquots of a pre-determined mass.

5.3 Results

Peptides targeting degron-binding sites within Cdc20 or the Apc8 subunit of the APC/C were designed using consensus degron sequences obtained from APC/C substrate proteins (Figure 5.4) [165, 166, 224]. To aid the design process, X-ray crystal structure data of proteins or small molecules bound to the degron-binding sites of Cdc20 [193, 195, 225] and Cdh1 [165] were also used. Sequences, structures and analytical data relating to all synthesised peptides are contained within Appendix B and are summarised later in Table 5.1.

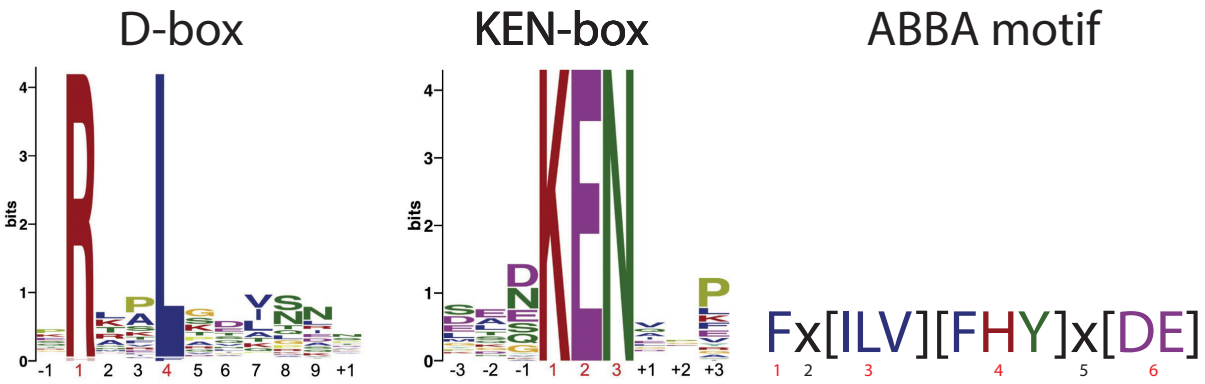


Figure 5.4: Consensus binding motifs derived from 68 APC/C substrates for the D-box and KEN-box motifs (adapted from He J. *et al.* 2013), [165]. Consensus ABBA motif adapted from Di Fiore *et al.* 2015, [166]. Numbers labelled in red describe amino acids necessary for co-activator binding.

5.3.1 D-box peptide design

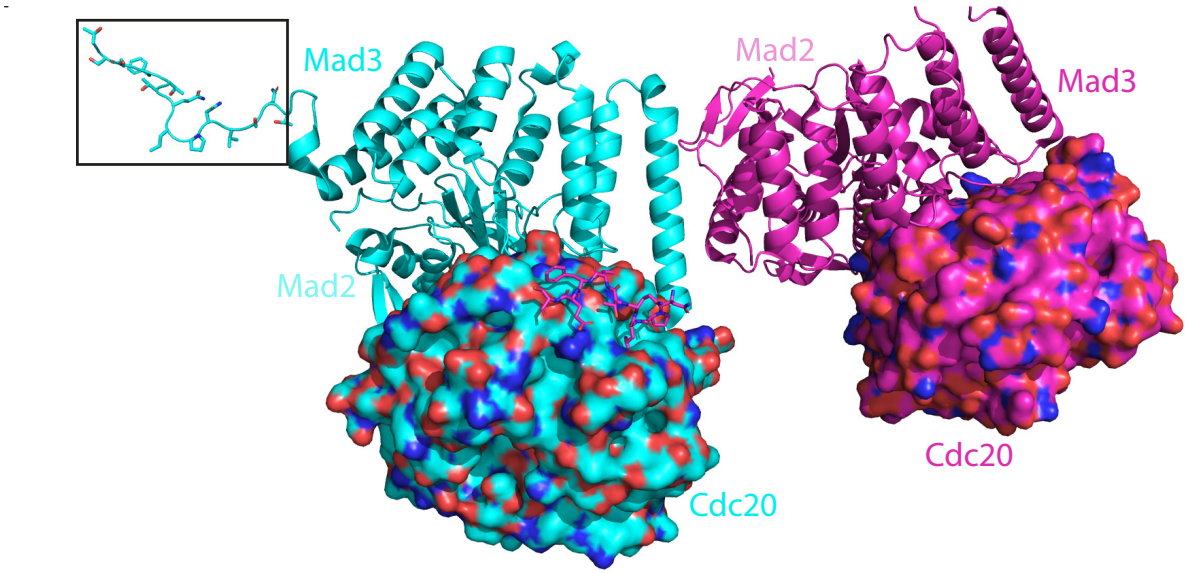
Sequences of the D-box peptide

The D-box peptide is comprised of nine amino acids, [165] and is the most studied of all motifs in the literature with 119 D-box degrons experimentally validated across all organisms [168, 169]. For all peptides the canonical nine-residue motif was used. The mitotic checkpoint complex (MCC) is a multi-subunit complex that binds to the APC/C to

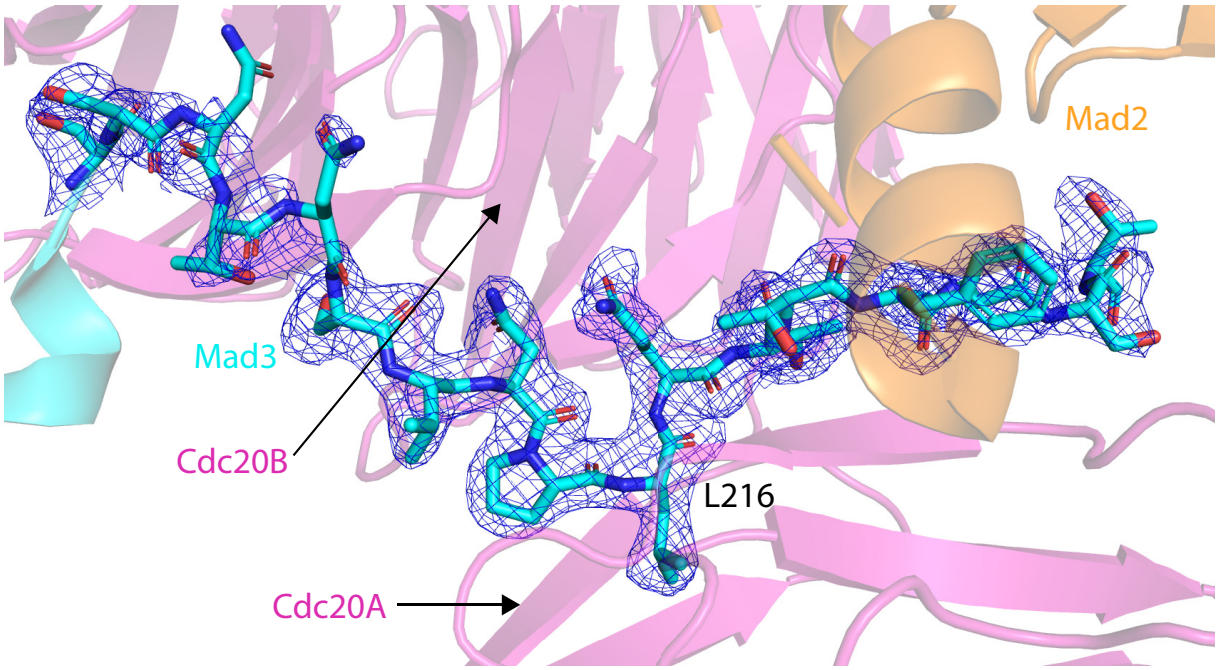
prevent premature anaphase onset, if chromosomes have not fully attached to the mitotic spindle [173, 224, 225]. The structure of fission yeast MCC (PDB:4AEZ) was determined at 2.3 Å resolution as a heterotrimer of Cdc20, Mad2 and Mad3. Mad3 interacts with the D-box and KEN-box binding sites of Cdc20 simultaneously [225]. Figure 5.5a illustrates how the C-terminus of Mad3 from one MCC heterotrimer (magenta) binds to the D-box-binding site of Cdc20 contained within the other MCC heterotrimer (cyan, surface representation) in the same asymmetric unit. This inter-MCC interaction is believed to be a crystallographic artefact of packing, and Chao *et al.* detail how the Mad3 C-terminus provides a D-box peptide mimetic [225]. Nonetheless, this structure provided the first atomistic detail of how a D-box peptide may interact with the $Cdc20^{WD40}$ domain. The C-terminal residues of Mad3 appear to be better resolved in one heterotrimer (cyan) and Coot was used to further examine the local architecture, contoured at 1.5σ in the 2Fo-Fc map (Figure 5.5b). The electron density of amino acids C-terminal to Leu216 within the Mad3 D-box peptide mimic are well defined. However, it is unclear whether this is a true representation of the interactions that may occur in natural D-box degrons, or whether it is a crystal packing artefact, in which the residues appear to be wedged tightly between the $Cdc20^{WD40}$ domain and Mad2 of the adjacent asymmetric unit.

In addition to the MCC, the structure of the *S. cerevisiae* Cdh1-APC/C modulator 1 (Acm1) complex was determined at 2.9 Å resolution (PDB:4BH6) [165]. Acm1 is a natural inhibitor of Cdh1, that interacts through three peptide motifs resembling ABBA, KEN-box and D-box degrons. This unique structure supported the architecture of D-box degron binding observed in the MCC. As can be seen in Figure 5.6, the Cdh1-Acm1 structure provided additional evidence for the binding of the canonical arginine at position 1 of the degron, in addition to a plausible binding pose for a leucine at position 7 of the degron (numbering according to Figure 5.4). However, the binding position of L125 is non-standard for non-polar aliphatic amino acids: the hydrophobic side chain is solvent exposed and appears to stack against the aliphatic atoms of D256 of Cdh1. Although an interaction between the two side chains would not normally be expected, the model is well supported by the electron density well and densities for other possible rotamers were not observed. Since L125 is not surrounded by any neighbouring proteins within the crystal lattice, this interaction is unlikely to be an artefact of crystal packing.

In addition to the atomistic detail provided by the above crystal structures, a D-box degron derived from the Hsl1 protein sequence, termed the 'super D-box', was reported to be the most optimal binding sequence for the D-box degron binding site [164, 168, 169, 226]. This made the Hsl1-derived peptide an ideal starting point for peptide D1 (Ac-GRAALSDITN-NH_2) in addition to the consensus sequence, peptide D2



(a)



(b)

Figure 5.5: Structural representations of the D-box binding site in fission yeast Cdc20 in the MCC (PDB:4AEZ [225]). (a) C-terminus of Mad3 in the MCC heterotrimer (magenta, stick representation) mimics D-box peptide binding to Cdc20^{WD40} domain (cyan, surface representation). The black box highlights the area further explored with electron density maps. (b) The Mad3 pseudo D-box degron (cyan, stick representation) is bound to Cdc20A (magenta, cartoon representation) is wedged between Cdc20B (magenta, cartoon) and Mad2 (orange, cartoon) of the adjacent asymmetric unit. Electron density map (blue mesh) is contoured at 1.0σ in the 2Fo-Fc map.

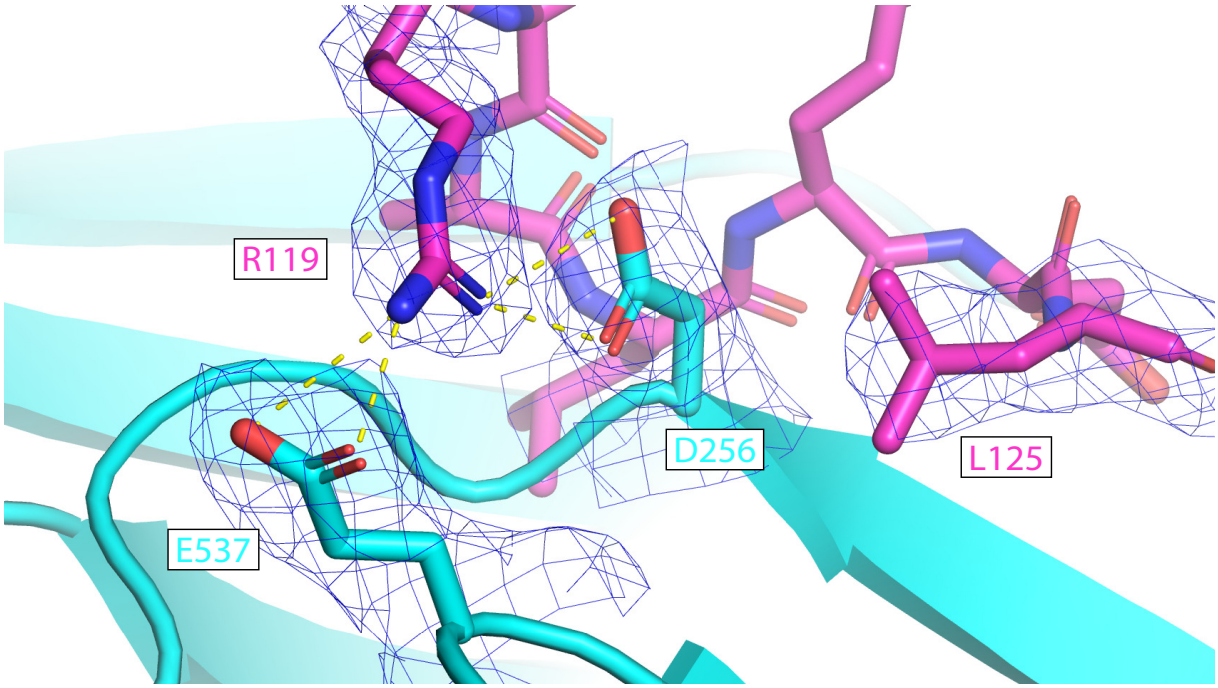


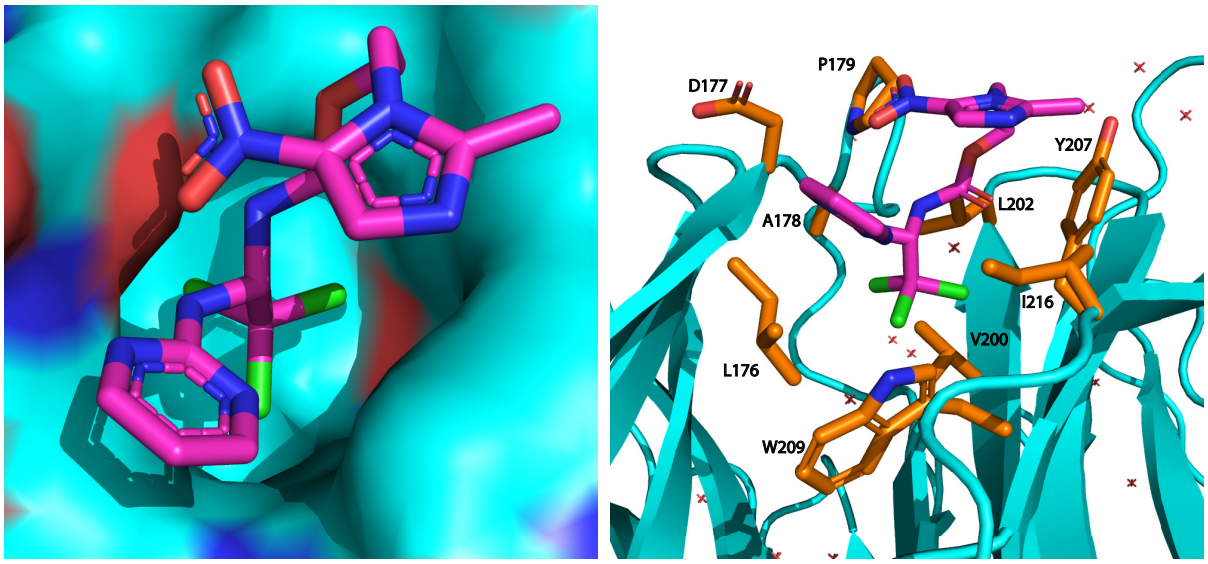
Figure 5.6: Structural analysis of the D-box binding site of Cdh1 bound to Acm1 (PDB:4BH6, ^{119}RIALKDL^{125}) [165]. R119 of Acm1 forms two hydrogen bonds (yellow, dashed lines) with Cdh1 with the carboxylic acid side chains of D256 and E537. Aliphatic side chain atoms of Acm1 L125 and Cdh1 D256 stack in a planar orientation. Electron density map (blue mesh) is contoured at 1.0σ in the 2Fo-Fc map.

(Ac-RLPLGDVSN-NH_2). However, due to the evidence of residue 2 remaining solvent exposed (observed in PBD:4AEZ [225] and PDB:4BH6 [165]), a large aliphatic amino acid such as leucine would likely be energetically unfavourable while also reducing the solubility. Therefore, leucine at position 2 was replaced by alanine in subsequently synthesised peptides, similar to the Hsl1-derived peptide sequence. In addition a negative control D-box peptide with the key binding residues removed was synthesised (D13: Ac-GDAAASDGTN-NH_2).

Exploring non-natural amino acids

The X-ray crystal structure of CdC20 bound to the small molecule Apcin was used for the design of further D-box peptides. Apcin interacts with Cdc20's D-box binding site, specifically docking into the pocket where the leucine sidechain at position 4 would dock (formed by L176, D177, A178, P179, V200, L202, Y207, W209 and I217, Figure 5.7b). The structure of Apcin bound to Cdc20 provides evidence that a larger chemical moiety compared to a leucine sidechain can bind this pocket, (Figure 5.7). To further gauge what size moiety may bind this pocket, the bond lengths of the CCl_3 moiety of Apcin were measured. Whereas the typical bond length of an aliphatic chain C-C bond is $\approx$

1.52 Å, the C-Cl bond of Apcin is slightly longer, at 1.73 Å. The total diameter and thus volume of the CCl_3 moiety fills more space within the degron-binding pocket. Additionally, halogenated ligands are typically more hydrophobic and 'drug-like' compared to their carbon-substituted counterparts. Given this analysis, non-natural amino acids with bulkier side chains than leucine were considered with the aim to increase binding affinity (Figure 5.8), and thus peptides D6 (Ac-RAPcBuAlaGDISN-NH_2), D7(Ac-RAPC_3GDISN-NH_2) and D20 (Ac-RAPC_3SDITN-NH_2) were produced. The synthesis of Fmoc-protected (*S*)-2-amino-4,4,4- trichlorobutanoic acid (Figure 5.8d was tasked to an external company, as there were no literature reports of its synthesis. Unfortunately, over 10 months, the synthesis of this amino acid was unsuccessful, and therefore it was not possible to introduce it into D-box peptides at position 4.

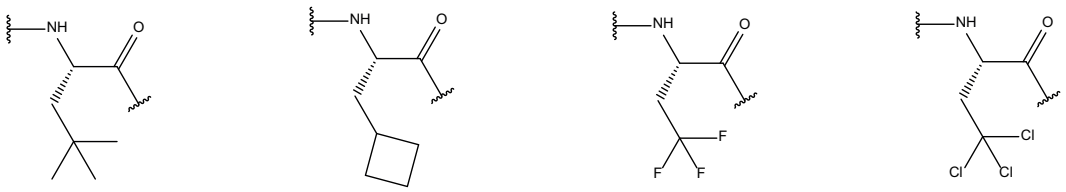


(a) **(b)**

Figure 5.7: Structural analysis of the D-box-binding site in human Cdc20 bound to Apcin, (PDB:4N14) [195]. (a) Apcin (stick representation, magenta) bound to human Cdc20 (surface representation, cyan) in the leucine-binding pocket of the D-box-binding site. (b) Cdc20 (cartoon representation, cyan) with amino acid side chains of leucine binding pocket forming residues illustrated (stick representation, orange).

Designing hydrocarbon peptide constraints

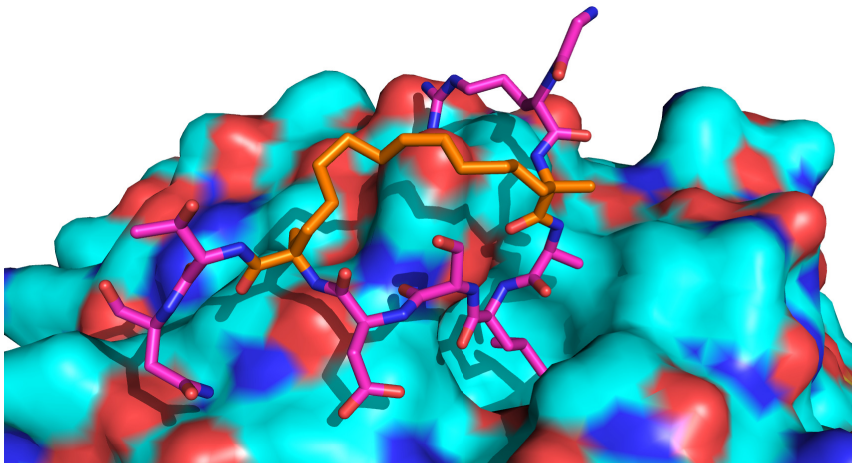
Hydrocarbon constraints of D-box peptides were considered under guidance from Dr. Yaw Sing Tan (Bioinformatics Institute, A*STAR, Singapore) and Dr. Fernando Ferrer (p53 laboratory, A*STAR, Singapore). Although many different chemistries exist for peptide stapling, only olefin-metathesised hydrocarbon staples were considered since Dr. Ferrer



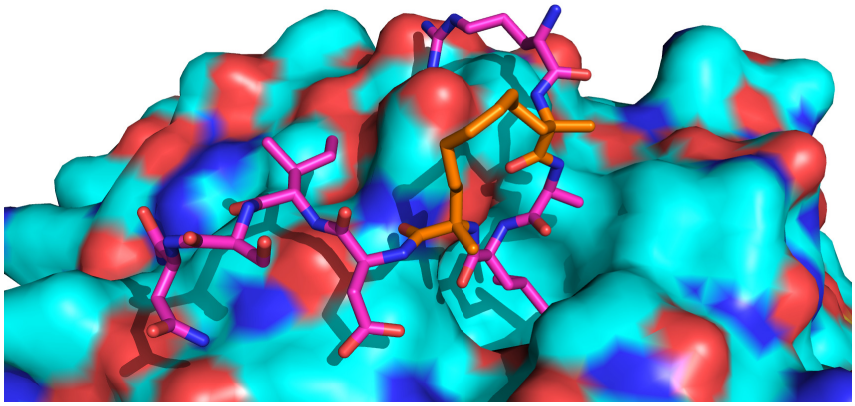
(a) C_3 **(b)** cBuAla **(c)** F_3 **(d)** Cl_3

Figure 5.8: Non-natural amino acids considered for position 4 in D-box peptides. (a) (*S*)-2-amino-4,4-dimethylpentanoic acid. (b) (*S*)-2-amino-3-cyclobutylpropanoic acid. (c) (*S*)-2-amino-4,4,4-trifluorobutanoic acid. (d) (*S*)-2-amino-4,4,4- trichlorobutanoic acid.

had extensive experience with this chemistry. Although stapling chemistries are more commonly used for the constraint of α-helical structures (peptides and peptoids), the constraining of loop structures [227], intrinsically disordered regions [228] and extended interfaces [102] has been explored also. Therefore it was thought to be reasonable to explore the constraint of D-box peptides. Due to the small number of residues in the peptide, only two stapling positions were explored (Figure 5.9). Peptide D9 (Ac-GRS_5ALR_8TN-NH_2) contains a longer hydrocarbon between the *S*-2-(4'-pentenyl) alanine (S_5) and *R*-2-(7'-octenyl) alanine (R_8) residues at positions 2 and 7 respectively, which are typically used in the stapling of α-helices. Although the aliphatic amino acid at position 7 was replaced, it was assumed that this loss could be overcome by confining the peptide towards a favourable binding conformation. Peptides D16 (Ac-RR_5AL$\mathbf{S_5}$DISN-NH_2), D17 (Ac-RR_5A$\mathbf{cBuAlaS_5}$DISN-NH_2) and D18 (Ac-RR_5A$\mathbf{C_3S_5}$DISN-NH_2) were designed using a shorter constraint with S_5 and S_5 amino acids at positions 2 and 5, respectively. This staple position was considered to be more optimal compared the positions used in peptide D9, since the aliphatic residue at position 7 in maintained. The distances between the Cα atoms of residues 2 and 5 are approximately 5.4 Å, which is akin to the i, $i + 3$ stapling of α-helices [229], although the geometries are very different. Despite an unproblematic synthesis, the RCM efficiency was very low for all three peptides. Kim *et al.* [229] also described poor RCM efficiency of the i, $i + 3$ stapled α-helices. Despite this, semi-preparative HPLC purification was attempted, however the samples were too low in purity and yield.



(a) D9



(b) D16

Figure 5.9: Models of stapling positions in D-box peptides. (a) Model of peptide D9 bound to Cdh1. (b) Model of peptide D16 bound to Cdh1. Models were built using the PyMOL builder function. All models were constructed using PDB:4BH6, using base coordinates for the Cdh1 (cyan, surface representation) and the Acm1 D-box peptide. D-box peptide shown as stick representation in magenta, with hydrocarbon staple amino acids coloured orange.

5.3.2 KEN-box peptide design

Unlike other WD40 or similar β-propeller domain proteins that bind their partners at the bottom-inner groove of the donut-like structure, the binding surface for the KEN-box is on the opposite face of the WD40 domain. This results in a very flat surface without any pockets to drive a small molecule drug discovery effort. Binding of the KEN-box motif to Cdc20 and Cdh1 is entirely driven by polar interactions, illustrated in Figure 5.10 [193]. The amino acids K, E and N of the KEN-box motif adopt a 3_{10} helix, which can be observed across all X-ray crystal structures of KEN-box peptides bound to Cdc20 (PDB:4GGD, 4AEZ) [193, 225] and Cdh1 (PDB:4BH6) [165] (Figure 5.11).

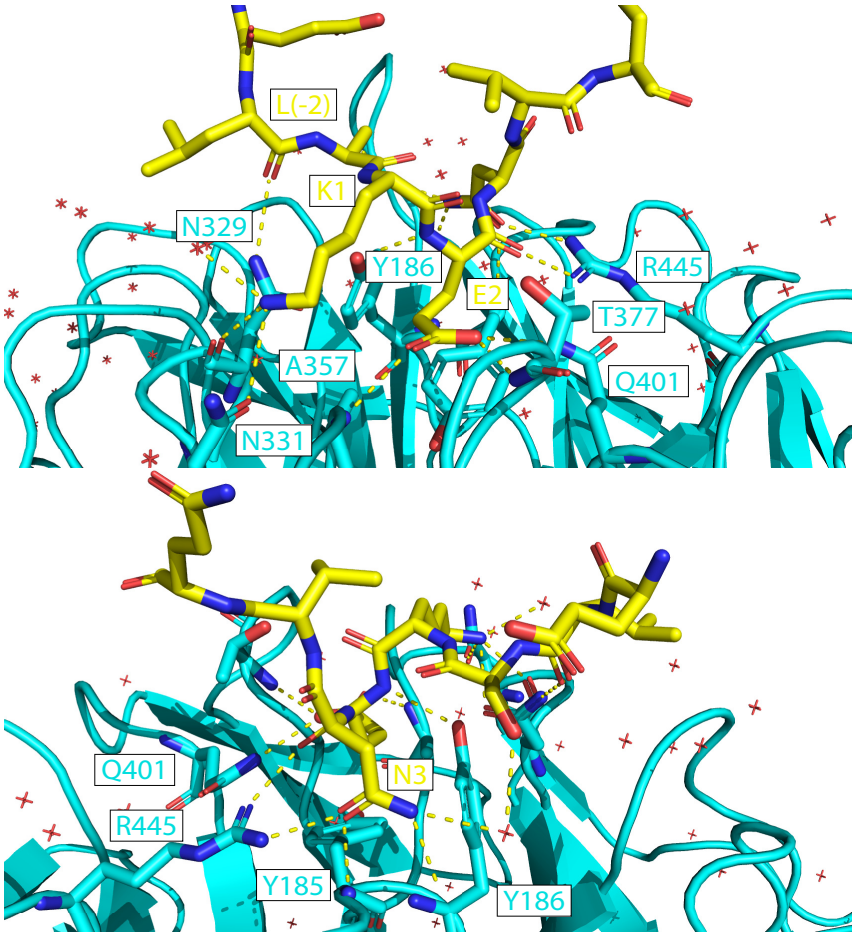
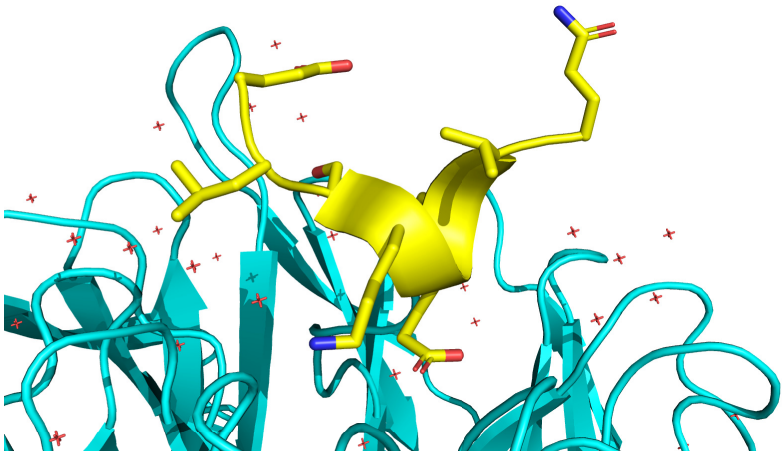
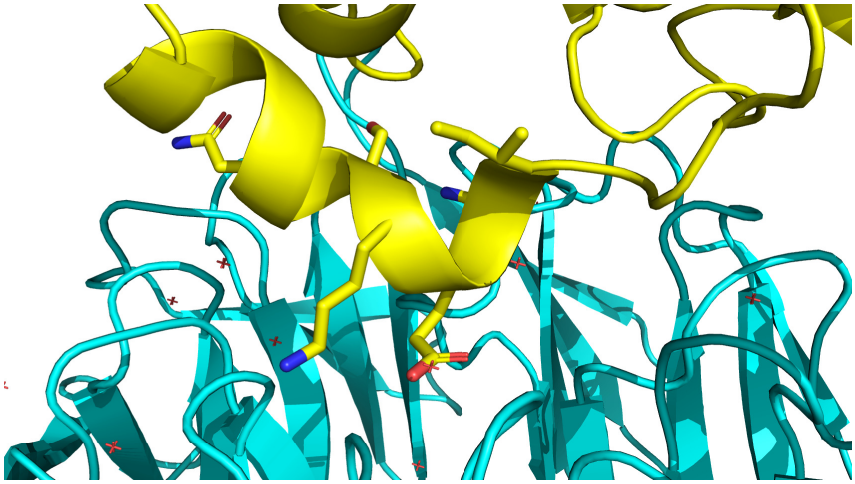


Figure 5.10: KEN-box peptide motif interactions with human Cdc20, PDB:4GGD. Polar interactions between the BubR1-derived KEN-box peptide (yellow, stick representation) and Cdc20 (cyan, cartoon and stick representation). Hydrogen bonds (dashed yellow lines) are formed between several side chain and backbone atoms between both molecules. Images are approximately a 180° rotation around the z-axis to show all intermolecular hydrogen bonds.

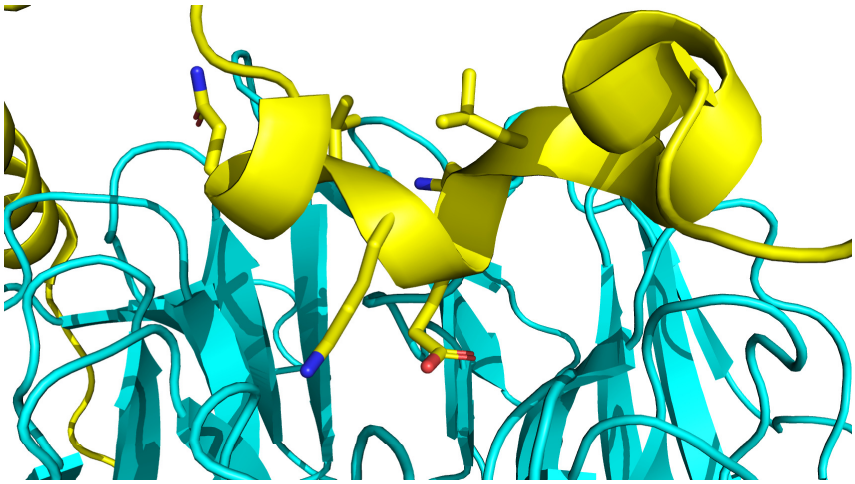
The first KEN-box peptide motif synthesised was based on the BubR1 sequence bound to Cdc20 from by Tian et. al from BubR1 [193]. However, the peptide used in their co-crystallisation experiment was significantly longer (CDEWELSKENVQPLRQGRIM-STLQ) compared to the residues observed. Therefore, a minimal motif (peptide K1, Ac-ELSKENV-NH_2) was designed based on the resolved residues in the hope that this motif would be sufficient for binding Cdc20. However, it is unlikely that the 3_{10} helix appearing in all KEN-box structures would be energetically favourable in the context of a minimal peptide. Therefore, hydrocarbon-constrained peptides were also designed and synthesised in an attempt to overcome this potential energetic barrier. Constraints were designed at positions 2 and 7, relative to the K1 peptide sequence and illustrative models of the constrained KEN-box peptides are shown in Figure 5.12. The synthesis of the hydrocarbon-constrained peptides, K7 (Ac-ER_5SKENS_5-NH_2) and K8 (Ac-ES_5SKENS_5-NH_2) with *R*-2-(4'-pentenyl) alanine (R_5)-S_5 and S_5-S_5 amino acids, respectively did not appear problematic, as evident from masses corresponding to the non-stapled peptide. However, similar to observations in D16, D17 and D18, the RCM reaction did not go to completion. Due to time constraints, further stapling chemistries and positions were not explored. A negative control KEN-box peptide with the key residues substituted for alanine was synthesised (K9: Ac-ELSAAAV-NH_2), however due to its insolubility in H_2O/MeCN, we were unable to purify this peptide.



(a) BubR1-Cdc20

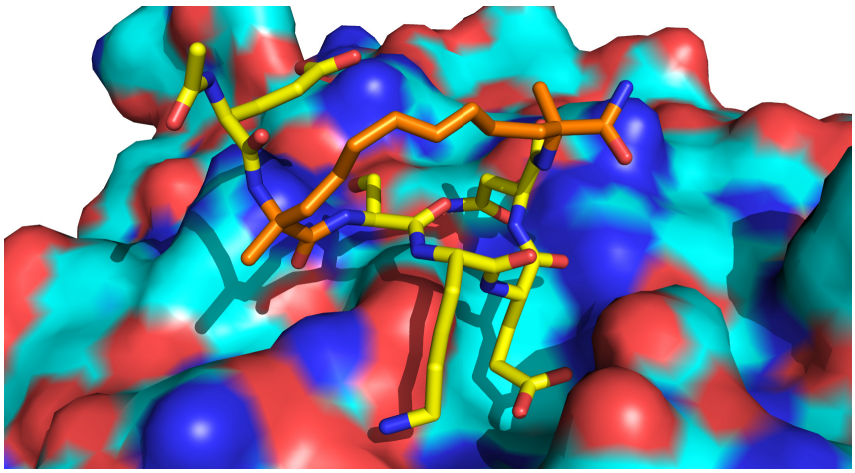


(b) *S. pombe* MCC

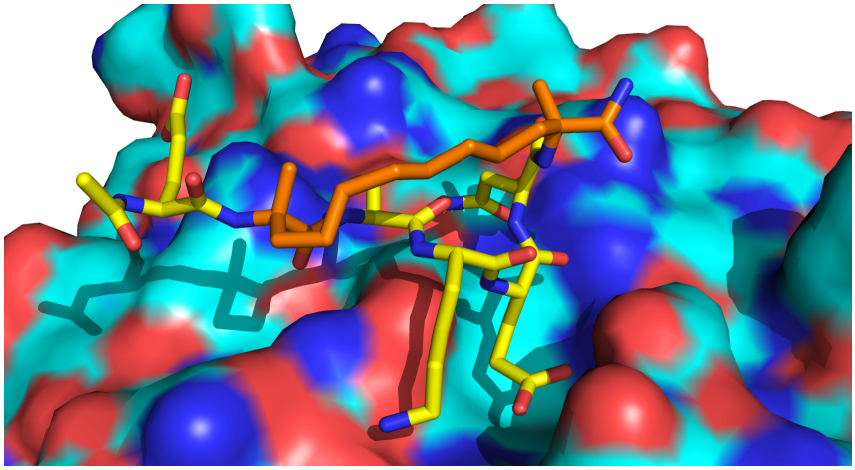


(c) Acm1-Cdh1

Figure 5.11: The KEN-box forms a strained α-helix turn motifs from three X-ray crystal structures of (a) BubR1 bound to Cdc20 (PDB:4GGD), (b) the *S. pombe* MCC (PDB:4AEZ), and (c) Acm1 bound to Cdh1 (PDB:4BH6). Proteins/peptides containing KEN-box motifs are shown in yellow with cartoon and stick representations. Cdc20/ Cdh1 are shown in cyan, cartoon representation.



(a) K7 model

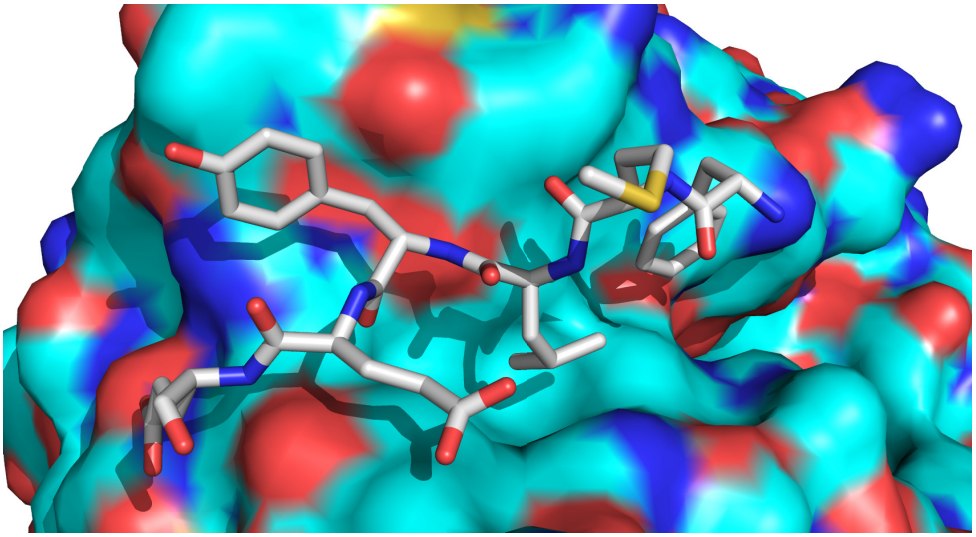


(b) K8 model

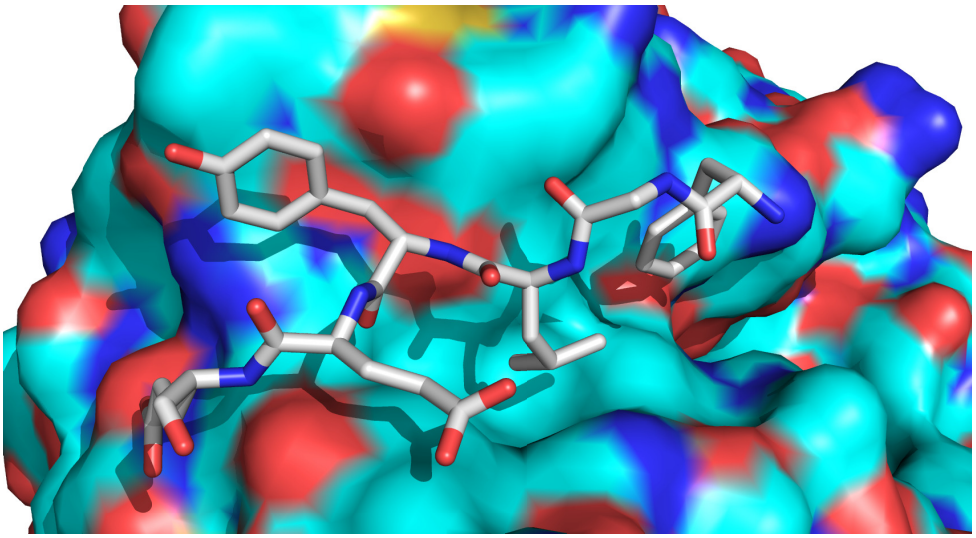
Figure 5.12: Models of stapled KEN-box peptides bound to Cdc20 based on coordinates from PDB:4GGD. (a) Model of peptide K7 with i, $i + 5$, R_5-S_5 hydrocarbon staple. (b) Model of peptide K8 with i, $i + 5$, S_5-S_5 hydrocarbon staple. KEN-box peptides are shown in yellow with hydrocarbon staple residues in orange. Cdc20 is shown in cyan as a surface representation.

5.3.3 ABBA motif peptide design

In contrast to the D-box and KEN-box degron motifs, there is only limited information on the ABBA motif as its discovery in *S. cerevisiae* Acm1 by Burton *et al.* is comparatively recent [230]. At present, only Acm1 [165, 230], Bub1 [166], Bub1B [151, 166, 167], Clb5 [231] and Cyclin A [166] are known to bind co-activators via ABBA motif degron sequences. As for the D-box and KEN-box peptide design, ABBA motif peptides were based on data from X-ray crystal structures [165] (Figure 5.13a and were kept minimal in length for ease of synthesis as well as to create a more 'drug-like' molecule. The ABBA motif relies on two hydrophobic interactions at positions 1 and 3, an aromatic ring at position 4 and a negatively charged residue at position 6 (numbering according to Figure 5.4). Peptide A1 (Ac-FGLYEE-NH_2, Figure 5.13b) was almost entirely based on the peptide motif from Acm1 [165, 230], apart from a methionine to glycine substitution at position 2. The methionine residue is solvent exposed and therefore unlikely to contribute to binding. Methionine also complicates the synthesis process, as peptide cleavage under acid conditions (TFA cleavage cocktail) yields would typically yield an oxidised methionine residue. Although protocols exist for the cleavage of methionine-containing synthetic peptides, it was decided to avoid the complication altogether. Since the ABBA motif adopts an extended conformation, hydrocarbon constraints would likely reduce the affinity and hence were not explored. A negative control peptide was also synthesised (A9: Ac-AGAAEE-NH_2), however due to insolubility issues this peptide could not be purified.



(a) Acm1 ABBA motif



(b) A1 model

Figure 5.13: (a) Binding pose of the Acm1 ABBA motif (residues 61-66) to *S. cerevisiae* Cdh1. PDB:4BH6 [165]. (b) Model of peptide A1 bound to *S. cerevisiae* Cdh1 based on PDB:4BH6.

5.3.4 C-box peptide design

In addition to targeting the degron-binding sites of Cdc20, peptides were designed to mimic the C-box motif contained within the N-terminal IDR of co-activators, which is partially responsible for binding to Apc8 of the APC/C [161, 232–234]. In addition to the C-box motif, co-activator association is mediated by the C-terminal Isoleucine-Arginine tail (IR-tail) at a structurally homologous binding site on Anaphase-promoting complex subunit 3 (Apc3) of the APC/C, (Figure 5.14) [162].

Peptide C1 (Ac-GDRYIPSR-NH_2), was designed using both human Cdc20 (GDRYIPHR) and Cdh1 C-box motifs (GRDFIPSR). Due to the smaller side chain, the serine in the Cdh1-bound peptide (Figure 5.15a) appeared to be more favourable than the histidine in the Cdc20-bound peptide, where intramolecular steric clashing could occur in alternative rotamers of the histidine side chain. Conversely, tyrosine at position 4 in the Cdc20-bound peptide (Figure 5.15b) appeared to be preferred over phenylalanine due to the potential ability to form an additional hydrogen bond. Peptide C5 (Ac-GDAAIPSA-NH_2) served as a negative control peptide with the key binding residues substituted for alanine residues.

Given that the C-box forms a tight loop comprising residues Y/F3 and R4, constraining this motif should be favourable for binding. Peptide C6 was designed with guidance from Dr. Yaw Sing Tan (Bioinformatics Institute, A*STAR, Singapore) who performed MD simulations to investigate an optimal hydrocarbon staple position and length. A

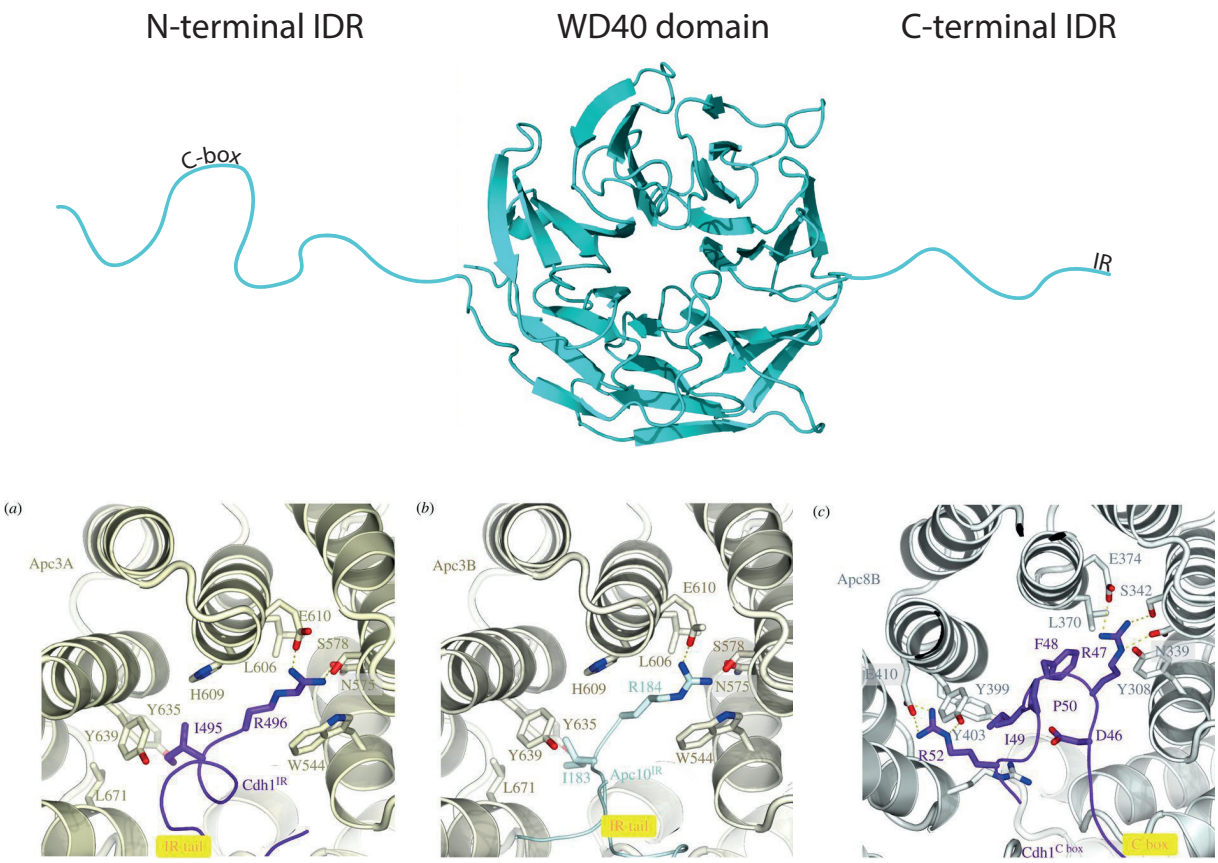
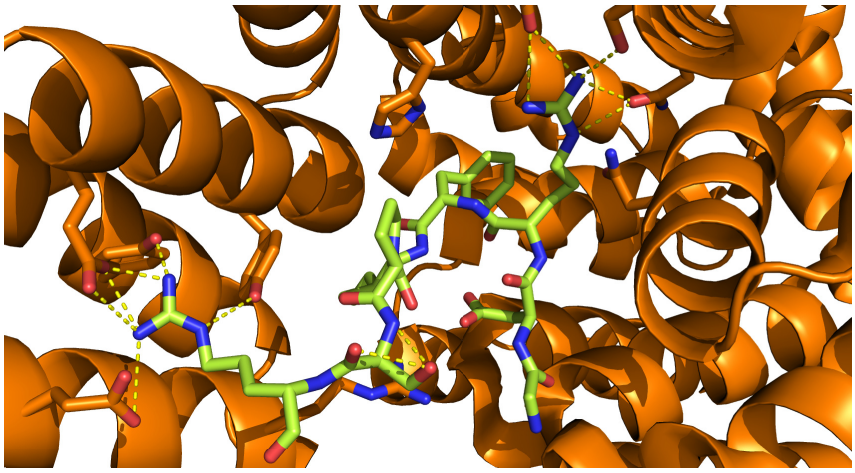


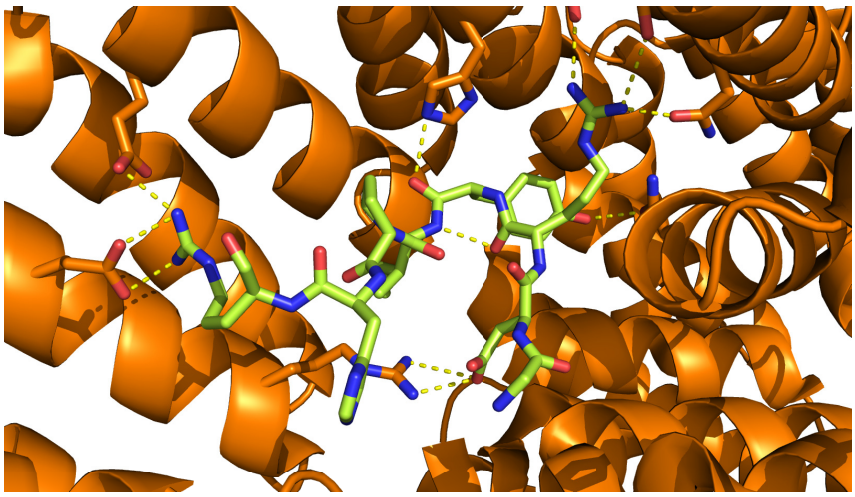
Figure 5.14: (Top) Schematic of co-activator domain topology with IDRs flanking the WD40 domain (Cdc20, PDB:4GGC [193]). (a) The Cdh1 IR-tail, (b) Apc10 IR-tail and (c) C-box motifs bind structurally homologous sites of Apc3 subunit A (Apc3A), Apc3 subunit B (Apc3B) and Apc8 subunit B (Apc8B), respectively. Panels (a, b and c) were adapted from Alfieri *et al.* [162].

model of the peptide is shown in Figure 5.15c. Peptide C6 (Ac-S_5DRYIPS_5R-NH_2) was then synthesised. Serendipitously, the semi-preparative HPLC provided sufficient separation of the *cis* and *trans* isomers produced during RCM. These peptides are labelled as 'C6e' (early isomer) and 'C6l' (late isomer) in Appendix B according to their analytical HPLC retention times. Further characterisation was not performed to identify the early or late peaks to the *cis* and *trans* isomers due to the difficulty of determining such isomers by proton nuclear magnetic resonance (^{1}H NMR) in the context of several hydrogen atoms within a peptide molecule.

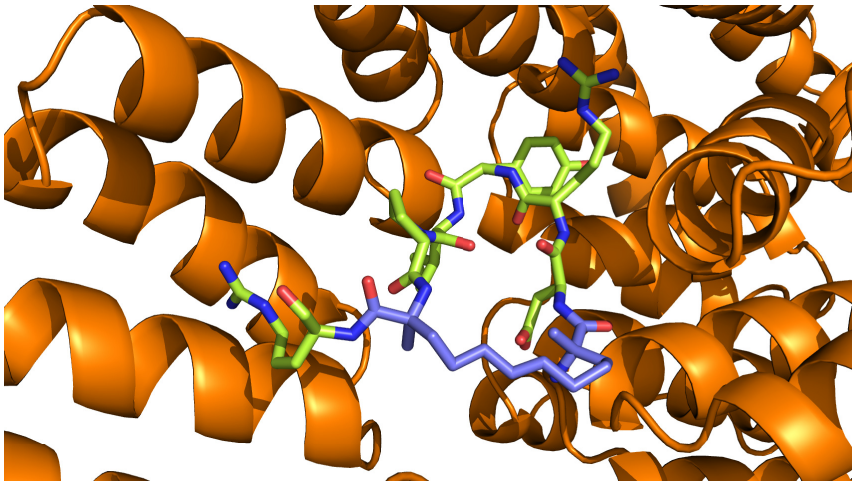
Since the binding sites of Apc8 and Apc3 are homologous, it was hypothesised that peptide C6 might also bind to Apc3. Alignment of the C6 peptide was performed against Cdc20 R499 of the IR-tail, using residues R3 and R8 (Figure 5.16).



(a) Cdh1 C-box

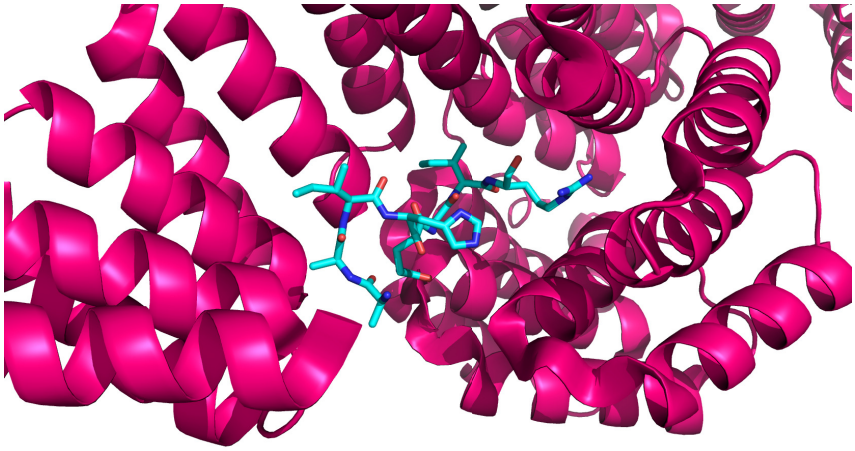


(b) Cdc20 C-box

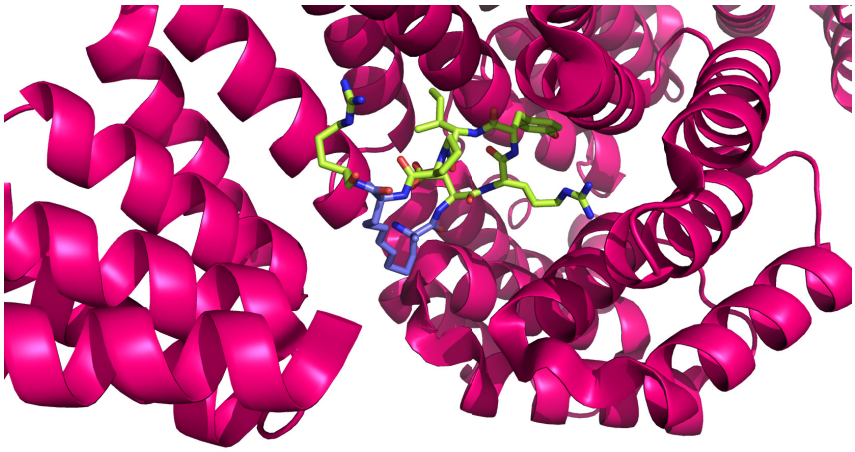


(c) C6 *trans* isomer model

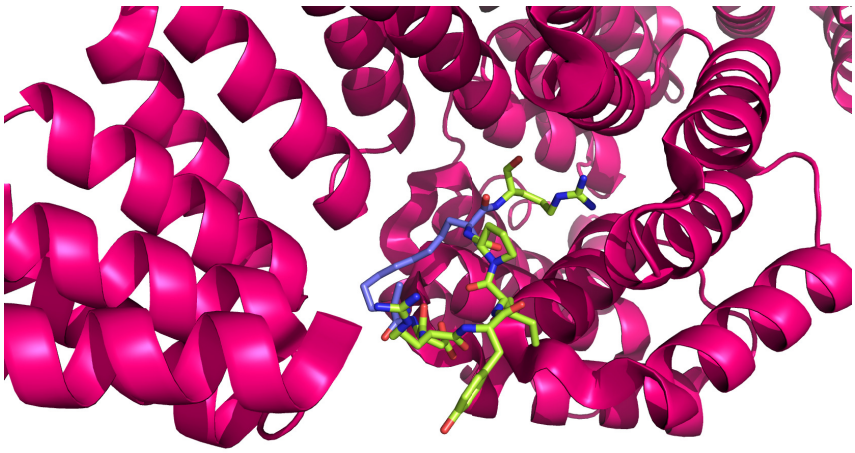
Figure 5.15: C-box peptides of (a) Cdh1 (PDB:4UI9) [150] and (b) Cdc20 (PDB:5LCW) [151] bound to the TPR superhelix structure of Apc8. (c) Model of C6 stapled peptide bound to Apc8. PDB:5LCW. C-box peptide motifs are shown in lime green and stick representation with hydrocarbon staple amino acids in blue. Apc8 is shown in orange, cartoon representation with stick representation for amino acids forming polar interactions. Polar interactions are indicated by yellow dashed lines.



(a) Cdc20 IR-tail



(b) R3 alignment



(c) R8 alignment

Figure 5.16: (a) Cryo-EM model of Cdc20-IR tail (cyan, stick representation) bound to the Apc3 subunit (magenta, cartoon representation). (b) Structural alignment of peptide C6, residue R3 to Cdc20 R499. (c) Structural alignment of peptide C6, residue R8 to Cdc20 R499. C6 peptides are shown in lime green, stick representation with hydrocarbon staples coloured in blue with Apc8 in magenta cartoon representation. (PDB:5LCW) [151].

5.4 Discussion

Four series of peptide inhibitors targeting the three degron-binding sites (degrons: D-box, KEN-box and ABBA motif) of APC/C co-activator Cdc20 and the C-box-binding groove of Apc8 were produced. Primarily, sequences of all peptides were designed based on X-ray crystal structures and Cryo-EM datasets as well as consensus sequence information. However, consensus sequences are only reliable for D-box and KEN-box degrons, as there are many of these degrons that have been identified in natural substrates. Contrastingly, only a limited number of ABBA motifs are currently reported in the literature (due to its recent discovery as a degron motif) and likewise little information exists for the C-box.

All peptide sequences are listed at the end of this chapter (Table 5.1). In SPPS, coupling of Fmoc-amino acids was initially performed for 1 hour per residue. This resulted in very low throughput, particularly at times when only one or two peptides were synthesised in parallel. Since the formation of the amide bond occurs very rapidly, an improvement to the solid phase synthesis was implemented whereby coupling reaction times were reduced to 10 minutes with no apparent loss of yield or generation of side products, which typically resulted in full peptide sequences synthesised in less than 2 days. The 0.1 mmol synthesis scale would normally result in around 50-100 mg of crude product. Following semi-preparative HPLC purification, 30-60 mg of purified peptide were typically obtained. Although no more than 1-2 mg of peptide were required for downstream biophysical analysis, the larger scale preparations were necessary in order accurately work with a known concentration of peptide. This was particularly important as amino acid analyses were previously conducted by Dr. Peter Sharrat (formerly PNAC facility, Biochemistry department, University of Cambridge). However since his retirement, this service was no longer available. Although commercial services exist, the amount of sample they required and cost of the service was prohibitive for concentration determination of peptides.

Due to the need to focus on a small number of peptides, the KEN-box and ABBA motif peptide series were relinquished early on, in favour of the D-box peptides. However, despite ceasing to work on these two motifs further, there is still plenty of scope for both peptides to optimise the length and amino acid sequences, incorporate non-natural amino acids similar to those explored in the D-box peptide and to further explore the constraint of the peptide. Indeed, one may also consider to make macrocyclic small molecules and even small molecules based on the ABBA motif, however these were considered beyond the scope and time constraints of the project.

For three out of the four peptide series, hydrocarbon constraints were explored. The lit-

erature is well defined in the design of hydrocarbon constraints in the context of α-helices. However, in the context of loop (D-box and C-box) or 3_{10}-helix (KEN-box) structures that comprise different geometries, no information exists. In addition, the short sequences that were designed also limit the number of stapling positions and so only a few hydrocarbon staples were designed and synthesised through a rational approach. Unfortunately, poor RCM efficiencies were observed in five out of seven hydrocarbon-contrained peptides. However, RCM reactions of peptides D9 and C6 went to completion. In addition, the isomers produced at the carbon-carbon double bond were efficiently separated in C6, which allows for the downstream biophysical characterisation of two distinct peptides. This was not possible for D9 as the crude yields after RCM were too low. The increased chromatographic separation of the two isomers appears unique to these loop-like structures, since the α-helical peptides do not separate efficiently. Despite the low success rate of producing stapled peptides, exploring this in more detail was outside the scope of the project due to time constraints. However, several different methods to constrain peptide structures including click-chemistry, lactam bridging and thioether formation among others (reviewed by Lau *et al.* [93]) exist and could be considered in future works.

Although a rational design approach is valuable and interesting for a PhD student, the throughput is low. Library-selection methods, such as phage and ribosome display, offer a significantly larger chemical space. However, at the beginning of this project and indeed till the end, generating purified Cdc20 was a limiting factor so these techniques were not explored.

Table 5.1: Summary of synthesised APC/C-targeting peptides. % purity is calculated using the integral of the peptide peak monitored by absorbance at 220 nm. Abbreviations of unnatural amino acids: C_3 - (*S*)-2-amino-4,4-dimethylpentanoic acid, cBuAla - (*S*)-2-amino-3-cyclobutylpropanoic acid, F_3 - (*S*)-2-amino-4,4,4-trifluorobutanoic acid, S_5 - *S*-2-(4'-pentenyl) alanine, R_8 - *R*-2-(7'-octenyl) alanine, R_5 - *R*-2-(4'-pentenyl) alanine. Full details of structures, HPLC chromatographs and LCMS data can be observed in Appendix B.

Name	Sequence	% purity	Synthesis outcome
D1	Ac-GRAALSDITN-NH_2	97.0	
D2	Ac-RLPLGDVSN-NH_2	92.4	
D3	Ac-RAPLGDVSN-NH_2	95.9	
D4	Ac-RAPLGDISN-NH_2	89.6	
D5	Ac-RAPLGDLSN-NH_2	89.9	
D6	Ac-RAPcBuAlaGDISN-NH_2	97.3	
D7	Ac-RAPC_3GDISN-NH_2	97.6	
D9	Ac-GRS_5ALR_8TN-NH_2	90.3	Mix of early and late isomers
D10	Ac-RAALGDISN-NH_2	99.7	
D11	Ac-RAAcBuAlaGDISN-NH_2	-	Insoluble; not purified
D12	Ac-RAAC_3GDISN-NH_2	98.6	
D13	Ac-GDAAASDGTN-NH_2	97.0	
D16	Ac-RR_5ALS_5DISN-NH_2	-	Poor RCM yield; impure
D17	Ac-RR_5AcBuAlaS_5DISN-NH_2	-	Poor RCM yield; impure
D18	Ac-RR_5AC_3S_5DISN-NH_2	-	Poor RCM yield; impure
D19	Ac-RAPLSDITN-NH_2	99.2	
D20	Ac-RAPC_3SDITN-NH_2	98.4	
D21	Ac-RAPF_3SDITN-NH_2	98.4	
K1	Ac-ELSKENV-NH_2	91.0	
K7	Ac-ER_5SKENS_5-NH_2	-	Poor RCM yield; impure
K8	Ac-ES_5SKENS_5-NH_2	-	Poor RCM yield; impure
K9	Ac-ELSAAAV-NH_2	-	Insoluble; not purified
A1	Ac-FGLYEE-NH_2	97.3	
A9	Ac-AGAAEE-NH_2	-	Insoluble; not purified
C1	Ac-GDRYIPSR-NH_2	98.2	
C5	Ac-GDAAIPSA-NH_2	84.9	
C6e	Ac-S_5DRYIPS_5R-NH_2	87.7	Early isomer
C6l	Ac-S_5DRYIPS_5R-NH_2	87.7	Late isomer

Chapter 6

Biophysical characterisation of APC/C^{Cdc20} peptide inhibitors

6.1 Introduction

Several methods for characterising ligand binding exist, although they differ in their sensitivity, sample requirements, throughput and measured parameters. Here, it was purified Cdc20^{WD40} that was the limiting factor. With this constraint in mind, methods such as isothermal titration calorimetry (ITC) and fluorescence polarisation (FP) were not considered feasible. Although FP experiments often require little protein, the amount is strongly dependent on the affinity of the fluorescent-labelled ligand for the protein. If ligand binding is weak, FP will require higher protein concentrations to obtain the 50-80% of the protein in the ligand-bound state at equilibrium, suitable for competition experiments. Therefore, differential scanning fluorimetry (DSF) and surface plasmon resonance (SPR) experiments were explored as alternative methods, since both have low protein requirements and are also high-throughput thereby allowing multiple peptide ligands to be tested in parallel.

6.2 Methods

6.2.1 Differential scanning fluorimetry

Experiments were conducted using a Roche Lightcycler 480 I or Biorad CFX Connect in 96-well plate format. Cdc20 (161-477) was prepared at a final concentration of 750 nM in 25 mM Tris-HCl, 150 mM NaCl, 1 mM $MgCl_2$, 5% (v/v) glycerol, 1 mM TCEP and 1% (v/v) DMSO, 5× Sypro Orange, pH 8.5 in the absence or presence of varying ligand concentrations in a final volume of 20 µL. Plates were sealed with a transparent film. Thermal melt curves were acquired using temperature ramps from 25°C to 95°C

at a rate of 0.03°C/sec and data were collected at a frequency of 20 points/°C for the Roche Lightcycler 480 I. Thermal melts on the Biorad CFX Connect were acquired using a temperature ramp from 25°C to 95°C with a step ramp of 0.5°C/step with a 20 second equilibration before reading each new temperature. An excitation wavelength of 483 ± 35 nm was used for the Sypro Orange dye and fluorescence was detected at an emission wavelength of 568 ± 20 nm.

Data analysis

Melting temperatures (T_m), were determined using the in-built 'T_m calling' function of the Roche LightCycler 480 I software. Briefly, the software calculates the negative derivative of the melt curve, where the minimum then corresponds to the T_m of the sample. Melting temperatures were derived separately for each well, averaged over triplicate well samples and are reported with the standard error of the mean (SEM). Data from the Biorad CFX Connect is automatically exported with the raw and negative derivative data.

6.2.2 Surface plasmon resonance

All experiments were conducted using a Biacore T200 instrument (GE healthcare) at AstraZeneca under the guidance of Dr. Christopher Stubbs.

Ligand immobilisation - Ni-NTA biosensor

Experiments on the Ni-NTA biosensor chip were conducted using His_6-TEV-Cdc20 (161-477) protein (see lane 'SPR SEC' in Figure 4.5a) in running buffer (10 mM HEPES, 150 mM NaCl, 0.1 mM TCEP, 0.05% (v/v) Tween 20 and 1% (v/v) DMSO) at 25°C. The His_6-TEV-Cdc20 (161-477) protein was diluted in running buffer to $\approx$ 20 µg/mL in 500 µL. The Ni-NTA biosensor was docked and primed three times in running buffer. Each flow cell was processed in succession. All steps were performed with a flow rate of 30 µL/min. First, flow cells were stripped of Ni^{2+} with 500 mM EDTA, followed by re-equilibration with running buffer. The flow cell was then recharged with 500 µM $NiSO_4$. 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide and *N*-hydroxysuccinimide were mixed and immediately flown over to activate the surface for covalent coupling. His_6-TEV-Cdc20 (161-477) was flown over Fc 2, 3 and 4 for increasing lengths of time, respectively, or with increasing RU capture targets between flow cells. All flow cells were blocked using 1 M Ethanolamine, pH 8.5. Non-covalently bound ligands were removed with an injection of 500 mM EDTA.

Ligand immobilisation - Streptavidin biosensor

The Streptavidin (SA) biosensor chip was prepared in a similar manner, but using biotinylated Cdc20 (161-477) protein (Figure 4.8) and the following alterations. The biotinylated Cdc20 (161-477) protein was diluted in running buffer 2-fold. Flow cells were washed three times with 50 mM NaOH, 1 M NaCl. Protein was then flowed at 5 µL/min for 5, 10 and 20 minutes over Fc 2, 3 and 4, respectively. All flow cells were then blocked with an injection of 50 µM amine-PEG_4-Biotin and detectors were then normalised with an injection of 70% (v/v) glycerol as per the manufacturers recommendation.

Analyte binding experiments

Compounds were diluted to a top concentrations of 100, 50 or 10 µM in running buffer maintaining 1 % (v/v) DMSO in the final solution, and used to created a 7-point 2-fold serial dilution in a 96-well plate. A running buffer blank was included as the first sample for each ligand. Vials for solvent correction were prepared using 0.2, 0.8, 1.4 and 2.0 % (v/v) DMSO and were run before and/or after all analyte binding experiments. Within a particular analyte sample, each series were run from low to high concentrations to reduce the chance of slow dissociation times effecting the next injection cycle. Injection cycles were set at 60 seconds of association and 60 or 120 seconds of dissociation at 30 µL/min. Between different analytes, a minimum of four running buffer injections were performed to ensure all analyte was removed prior to the injection of the next analyte.

Data analysis

The reference flow cell (Fc 1) was subtracted from all sensorgrams and the 0 µM blank belonging to the analyte dilution series to better account for any loss of ligand activity over the course of an experiment. Steady-state affinities were calculated from the concentration series of each analyte. Observed response units (RU) were calculated from a 5-second average, 12 seconds prior to the analyte injection end-point of each injection cycle and are relative to the RU at t = 0 sec of analyte injection. Since buffer blank subtractions were performed, all data were fitted using

$$RU_{analyte} = \frac{RU_{max} \times [analyte]}{(K_d + [analyte])}, \tag{6.1}$$

where $RU_{analyte}$ is the observed response units at a given analyte concentration at equilibrium, $[analyte]$ is the molar concentration of the analyte, K_d is the dissociation constant for the analyte and ligand interaction and RU_{max} is the theoretical maximum response of the analyte with a fixed ligand immobilised. RU_{max} is further defined by

$$RU_{max} = \frac{MW_{analyte} \times RU_{ligand} \times valency_{ligand}}{MW_{ligand}}, \tag{6.2}$$

where RU_{ligand} is the response units immobilised onto the biosensor chip, $valency_{ligand}$ is the stoichiometry of the interaction, $MW_{analyte}$ is the molecular weight of the analyte and MW_{ligand} is the molecular weight of the ligand. The percentage of functional ligand immobilised onto the biosensor chip, $\%\ Ligand_{functional}$, was calculated using

$$\%\ Ligand_{functional} = \frac{RU_{max}}{RU_{ligand}} \times \frac{MW_{ligand}}{MW_{analyte}} \times 100. \tag{6.3}$$

All values are reported as the mean and standard error calculated from the fits of reference and blank subtracted flow cells, Fc 2-1, Fc3-1 and Fc 4-1. Since RU artefacts often occur upon valve switching from running buffer to sample injection, and vice-versa, data at the beginning of these valve transitions are excluded from the raw data. These data are outside of the raw sensorgram data used for calculation of steady-state affinity and thus their exclusion does not affect the calculation of binding affinities as described above.

6.2.3 Cellular Thermal Shift Assays

CETSA of endogenous protein

HEK 293T cells were cultured in DMEM (Gibco) + 10% FBS (Sigma Aldrich) at 37°C, 5% CO_2, in a humidified environment. Cells were cultured in multiple T75 flasks to generate sufficient quantities of cell lysates for downstream applications. Cells were grown to confluency, washed in PBS and trypsinised from the culture vessel for 5 minutes at 37°C, before quenching with medium. Cell suspensions were pooled and centrifuged at 500 × g for 3 minutes to pellet the cells. Cell pellets were washed twice in PBS with repeated centrifugation. The washed and pelleted cells were resuspended in NET-2 buffer (50 mM Tris-HCl, 150 mM NaCl, 1 mM $MgCl_2$, 0.5 % (v/v) NP-40, 0.5 mM DTT, 2 mM $NaVO_3$, 5 mM NaF, 1 mM PMSF and 1 × SigmaFAST EDTA-free protease inhibitor cocktail (Sigma Aldrich), pH 7.4). The lysate was incubated at 4°C for 1 hour on a rotating wheel to lyse. The lysate was clarified by centrifugation at 21,000 × g, 4°C for 20 minutes and the supernatant was transferred to a fresh microfuge tube. The total protein concentration was determined as described in Section 2.3.1, yielding 2.27 mg/mL. Lysates were aliquoted in 100 µL, flash frozen in liquid N_2 and stored at -80°C.

For each experimental condition, a single 100 µL aliquot was thawed on ice. Apcin, TAME or DMSO were added to a final concentration of 100 µM compound and 1% DMSO, samples were mixed immediately by vortexing and returned to ice to incubate for 1 hour. All subsequent steps were performed on PCR machines. Incubated lysates were aliquoted into 11 PCR strip tubes at 4°C. Each strip tube was heated at a different temperature between 40 and 60°C, in 2°C steps for 3 minutes before returning to 4°C. Tubes were then centrifuged at 20,000 × g, 4°C for 10 minutes to pellet precipitated proteins. The super-

natant was added to an equal volume of 2× SDS PAGE loading buffer and samples were loaded onto an 8% SDS PAGE gel cast using ProtoGel buffers and acrylamide (National diagnostics), with Spectra^{TM} Multicolour Broad Range protein ladder (Thermo Fisher Scientific) in both outside lanes. Gels were run in TGS buffer at 60 V for 30 minutes followed by 120 V for 90 minutes. Gels were then transferred onto 0.45 µm PVDF membrane in transfer buffer using a semi-dry transfer system (Pierce Power blotter, Thermo Fisher Scientific). PVDF membranes were then cut (if necessary) and incubated in western blot blocking buffer for 1 hour at room temperature on a rocking platform. The respective membranes were incubated with 1:1000 dilutions of anti-Cdc20 (D6C2Q, #14866, Cell Signalling Technologies), anti-Apc8 (D5O2D, #15100, Cell Signalling Technologies) or anti-Apc3 (D3I1V, #12530, Cell Signalling Technologies) Rabbit monoclonal antibodies (mAbs) in western blot blocking buffer overnight at 4°C. The following day, primary antibodies were decanted and membranes were washed 3 × in PBST. Blots were incubated with a 1:10000 dilution of Swine anti-Rabbit HRP-linked polyclonal antibody (DAKO) in western blot blocking buffer for 1 hour at room temperature on a rocking platform. Secondary antibody was removed and membranes were washed three times with PBST. Amersham ECL western blotting substrate was prepared according to the manufacturers instructions and incubated with the blot for 5 minutes. Membranes were imaged using a BioSpectrum® 810 imaging system (UVP) with a cumulative integration of data over time. A white light image was also taken for reference. Data were analysed by preparing a box of fixed dimension around each band and calculating the integral of the signal peak using ImageJ [235]. The integrands were normalised relative to the highest observed signal in a thermal melt curve, plotted in GraphPad Prism 7.0 and fitted using a Boltzman equation to detemine the apparent melting temperature (T_m) [236]:

$$Y = Y_{min} + \frac{Y_{max} - Y_{min}}{1 + e^{(\frac{T_m - X}{slope})}} \tag{6.4}$$

where Y is the normalised fraction of protein at temperature X. Y_{min} and Y_{max} are the minimum and maximum plateaus of normalised signal, and *slope* is the gradient of the melt transition. The midpoint of the transition defines apparent melting temperature, T_m.

CETSA of exogenous protein

High-throughput CETSA experiments with exogenous proteins were based on protocols described by N. Martinez *et al.* [237]. HEK 293T cells were cultured in DMEM (Gibco) + 10% FBS (Sigma Aldrich) at 37°C, 5% CO_2, in a humidified environment. Cells were seeded on day 1 to reach 70% confluency on the following day. Transfection reagents were prepared by mixing 10 µg of plasmid DNA (see Table 6.1 and Appendix A for construct information) with 30 µL of Lipofectamine 2000 (Thermo Fisher Scientific) in 1 mL of OptiMEM (Gibco). The transfection complexes were incubated at room temperature for 5 minutes prior to their addition to the cell culture flask. Cell cultures were incubated for a further 48 hours, before washing in PBS and trypsinisation. Cell suspensions were washed and centrifuged as described above, and were resuspended in lysis buffer (PBS (Gibco), 5 mM NaF, 2 mM $NaVO_3$ and 1 × SigmaFAST EDTA-free protease inhibitor cocktail (Sigma Aldrich). Cells were lysed by freeze-thawing in liquid N_2 and the lysates were clarified by centrifugation at 21,000 × g, 4°C for 20 minutes. Supernatants were transferred to a fresh microfuge tube and the total protein concentration was determined in each lysate. Supernatants were aliquoted in 100 µL aliquots, flash frozen in liquid N_2 and stored at -80°C.

After initial pilot experiments, only the Cdc20 (1-499)-HiBiT and Apc8 (1-591)-HiBiT transfected lysate was carried forward to analyse ligand binding. Cdc20 (1-499)-HiBiT and Apc8 (1-591)-HiBiT lysates were diluted to 0.2 mg/mL and 0.25 mg/mL, respectively in lysis buffer and aliquoted into separate tubes according to the number of experimental samples. To probe binding to exogenous Cdc20, Apcin, D21, D20, D19, D7 or DMSO (1 µL of 10 mM compound in 100% DMSO) were added to a final concentration of 100 µM compound, 1% DMSO. To probe binding to Apc8, TAME, C1, C5, C6e, C6l or DMSO (1 µL of 10 mM compound in 100% DMSO) were added to a final concentration of 100 µM compound, 1% DMSO Tubes were mixed immediately by vortexing and returned to ice to incubate for 1 hour. Thermal melts were conducted as described above. The heated lysate samples were aliquoted (2 µL/well) into an AlphaPlate light-grey 384-well plate in triplicate at 4°C. Lysis buffer and a non-transfected cell lysate control were included as negative control samples. The Nano-Glo®HiBiT lytic detection system (Promega) was prepared according to the manufacturers' instructions and 10 µL were added to each well. The plate was incubated on a shaking platform (1000 rpm) to ensure sample homogeneity. The plate was then read in a CLARIOStar microplate reader (BMG labtech) with the focal height and gain adjusted to 10.5 cm and 2000, respectively, and with the detector set to read the luminescence signal at 460 ± 80 nm. Data were normalised with respect to the first temperature of the thermal melt curve (40°C). Normalised data were then plotted in GraphPad Prism 7.0 and fitted using Equation 6.4.

Table 6.1: Cdc20, Apc3 and Apc8, HiBiT-tagged constructs used for transfection into HEK 293T cells for CETSA experiments with D-box and C-box-targeting ligands.

Construct	Tag terminus	MW (Da)
HiBiT-Cdc20 (1-499)	N	56444.6
Cdc20 (1-499)-HiBiT	C	56588.8
HiBiT-Apc3 (1-824)	N	93590.3
Apc3-HiBiT (1-824)	C	93459.1
HiBiT-Apc8 (1-591)	N	70007.6
Apc8-HiBiT (1-591)	C	69876.4

6.2.4 *In vitro* ubiquitination assays

In vitro ubiquitination experiments were performed by Dr. Thomas Tischer in the Barford laboratory at the MRC LMB, Cambridge, UK. Experiments were performed using APC/C and Cdc20 purified from insect cells as described by Zhang et al. [28]. 60 nM APC/C, 30 nM Cdc20, 90 nM UBA1, 300 nM UbcH10, 300 nM Ube2S, 35 μM ubiquitin, 1 μM cyclin B1, 5 mM ATP, 10 mM $MgCl_2$ were mixed in a buffer containing 40 mM HEPES, 80 mM NaCl, 0.6 mM DTT, pH 8.0. Reactions were incubated for 30 minutes at 23°C in the presence of ligands or DMSO as the vehicle control. Reactions were stopped by the addition of one volume of 2× concentrated NuPAGE LDS loading buffer (Invitrogen). All antibodies used and western blotting procedures are consistent with the previously published data from the Barford laboratory [28].

6.2.5 X-ray crystallography

Crystallisation

Cdc20 (161-477) from multiple protein purification attempts were combined for crystallisation trials. Proteins were concentrated using an Amicon Ultra 0.5 mL 10 K MWCO centrifugal filters (Merck) at 4°C, 14,000 × g. Proteins were concentrated to ≈ 0.4 mg/mL before peptide D21 was added to a final concentration of 100 μM, 1% DMSO, to stabilise the protein during further concentration and eventually yielding 1.9 mg/mL. A custom crystallisation grid screen was designed around conditions described by Tian *et al.* [193] (Figure 6.1). The solutions were dispensed into a 96-well 2-drop MRC crystallisation plate (Swissci, Molecular Dimensions) using a Dragonfly crystal instrument (TTP Labtech) and mixed for 5 minutes using a MXone instrument (TTP Labtech). The Cdc20 (161-477)-D21 complex was then dispensed using a Mosquito HTS (TTP Labtech). Drops in 5% glycerol well solutions were prepared using the 0% glycerol counterpart well solution. For

example, drop A4 was prepared from 100 nL protein complex and 100 nL of well solution from well A1. The crystallisation plate was sealed with Crystal clear sealing tape (Hampton Research) and stored at 20°C in a Rock Imager (Formulatrix). Visible light and UV images were collected to monitor each drop during the incubation period.

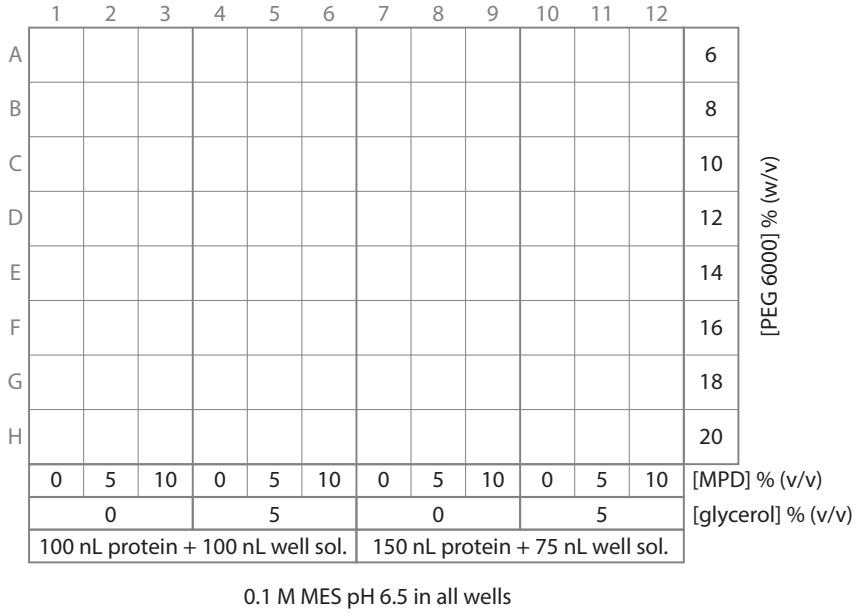


Figure 6.1: Setup of 96-well plate for crystallisation trials. All wells contain a fixed concentration of 0.1 M MES pH 6.5, varying concentrations of Polyethylene glycol (PEG) 6000 from 6-20 % (w/v), 0, 5 or 10% (v/v) 2-Methyl-2,4-pentanediol (MPD), 0 or 5% (v/v) glycerol and 1:1 or 2:1 ratios of protein to well solution in crystallisation drops.

Ligand soaking and crystal harvesting experiments

Ligand soaking experiments were performed by Dr. Marianne Schimpl at AstraZeneca using an adapted protocol from Sackton *et al.* [195]. Crystals were looped and washed through 3 × 0.5 µL drops of MPD-washout solution (0.1 M MES pH 6.5 and 20% PEG 6000) and were incubated in the final wash drop for 30 minutes at 20°C. Peptides were diluted to 2.5 mM (5% (v/v) DMSO) or 5 mM (10% (v/v) DMSO) in the MPD-washout solution. Crystals were passed through one drop (0.5 µL) of the peptide-containing solution and were incubated in a second drop (0.5 µL) of the peptide-containing solution for 4 hours at 20°C. For each peptide-soaked crystal, 3µL of cryo-protectant solution (0.1 M MES pH 6.5, 20% (w/v) PEG 6000 and 10% (v/v) glycerol) was mixed with the first peptide-containing soak solution. Crystals were looped and passed through the soak/cryo-protectant solution and were flash frozen in liquid N_2.

X-ray diffraction

Crystals were diffracted on beamline I04 at Diamond Light Source (Oxford, UK) by Dr. Jason Breed (AstraZeneca). The beam wavelength and size were fixed to 0.9795 Å and 32 µm × 20 µm, respectively. Crystals were manually centred and 1350 diffraction images were collected using an Eiger2 XE 16M detector (Dectris) over 270°, with 0.2° oscillations between each image (50 ms exposure time) at 100 Kelvin.

Structure determination

Diffraction images were processed using autoPROC-STARANISO [206, 238–242]. Criteria used in the determination of diffraction limits were set by a local $I/\sigma I \geq 1.20$. Structures were solved by molecular replacement using Phaser [209] with the Apo-Cdc20 (161-477) structure, PDB: 4GGC [193] followed by iterative model building in Coot [213] and refinement in Refmac5 [210] using 4.9% of reflections as an R_{free} test set. Restraint dictionary files for non-natural amino acids, (*S*)-2-amino-4,4-dimethylpentanoic acid (PDB ligand ID: 0JY) and (*S*)-2-amino-4,4,4-trifluorobutanoic acid (PDB ligand ID: 3EG) were generated using Grade v1.105 [243]. Ligands were built into the models after proteins and water molecules were refined and added.

Subsequent refinement of all structures were performed using BUSTER version 2.10.3 [211], with additional use of Local Similarity Structural Restraints (LSSR) for refinement the Cdc20-D7 complex [212]. Results pertaining to structural data are updated using newly refined datasets.

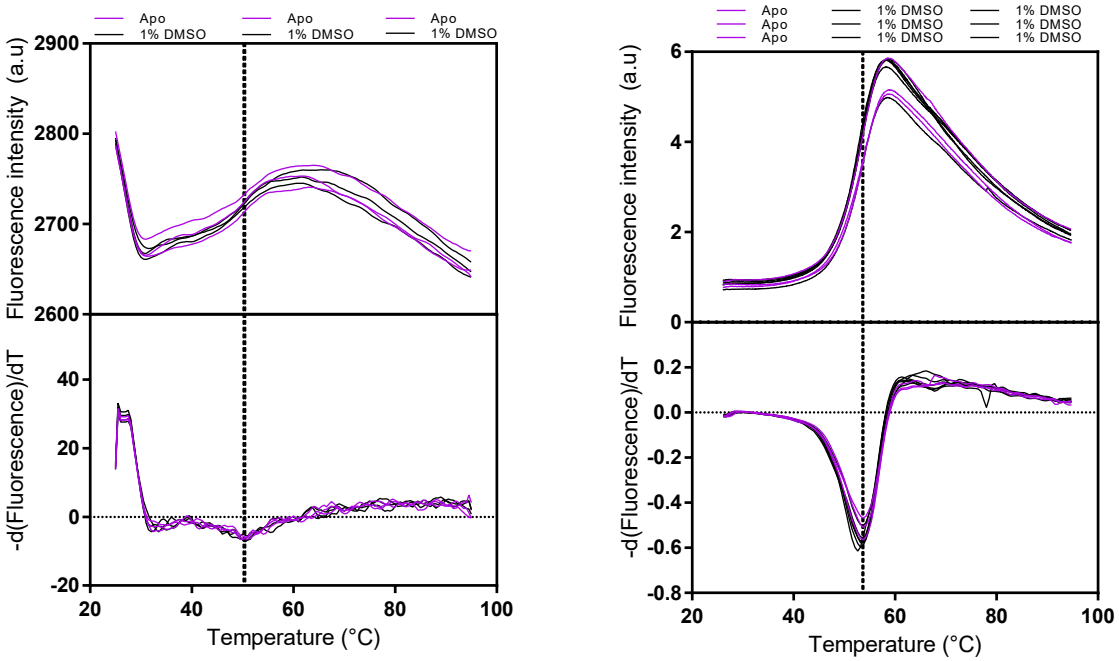
6.3 Results

6.3.1 Method development and initial screens of screens of degron peptides

Setting up DSF experiments

Differential scanning fluorimetry (DSF) was used to identify ligand-dependant stabilisation of recombinant Cdc20, in which the fluorescence intensity of the Sypro Orange dye molecule increases upon interacting with hydrophobic regions of a protein that are exposed during the unfolding of of Cdc20. Initial pilot experiments were performed at 1.5 μM of Cdc20 using the Biorad CFX Connect instrument, however, the data quality was not ideal. By serendipity around a similar time to developing this assay, our laboratory acquired a Roche Lightcycler 480 I, which can also be used for the same experiment. This allowed us to make a comparison of the data obtained from both instruments, where the results of Apo and vehicle control (1% DMSO) were used to make this comparison (Figure 6.2). Between the two machines, there are several noticeable differences. The first is the absolute T_m value for both Apo and vehicle control samples, which are 50.5°C and 50.2°C for the Biorad CFX connect and 53.8°C and 53.6°C, respectively. There is also more noise in the raw data and the density of collected data is approximately ten times less from the Biorad CFX connect compared to the Lightcycler 480 I, which has a direct impact on the calculation and peak height of the negative differential plot. This introduces a greater amount of uncertainty when determining the minimum. In addition, it was possible to use half the concentration of Cdc20 to 750 nM, and still retain high quality data. These differences may be due to multiple factors inherent to the individual instrument hardware and the differing protocols that can be used as a result of this. The Lightcycler 480 I is able to read all wells, simultaneously in a continuous manner, whereas the Biorad CFX Connect can only read single wells and thus must scan the entire plate at a fixed temperature. This results in longer runs and effectively slower ramp rate, which may alter the kinetics of protein unfolding and the rate of Sypro Orange dye association/ dissociation from the unfolded protein. In addition to the inherent hardware differences, it is also possible that the quality of proteins between batches could also be a causative factor of data quality. Unfortunately, given the limited amount of protein purified, it was not possible to compare this factor on one machine.

In the pilot experiment using the Biorad CFX connect, a small group of ligands including Apcin, D1, D9, A1 and K1 were also tested, in addition to the apo and vehicle controls (Figure 6.3). Data for D9, A1 and K1 were not re-tested using the Roche Lightcycler 480 I and thus may account for the poorer data quality comparatively to D1 and

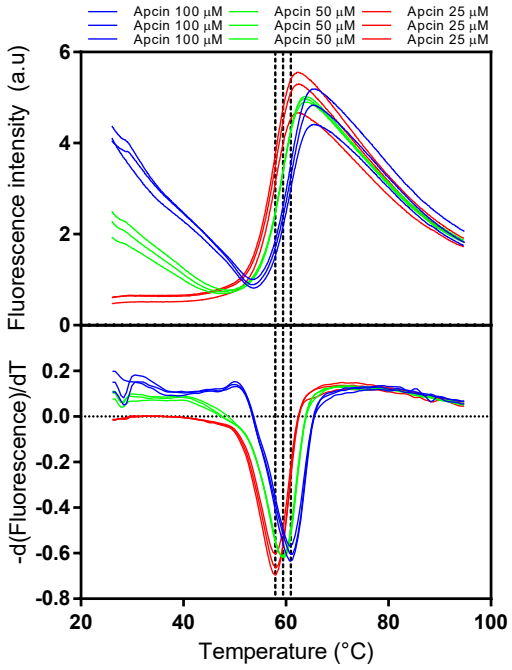


(a) Biorad CFX connect

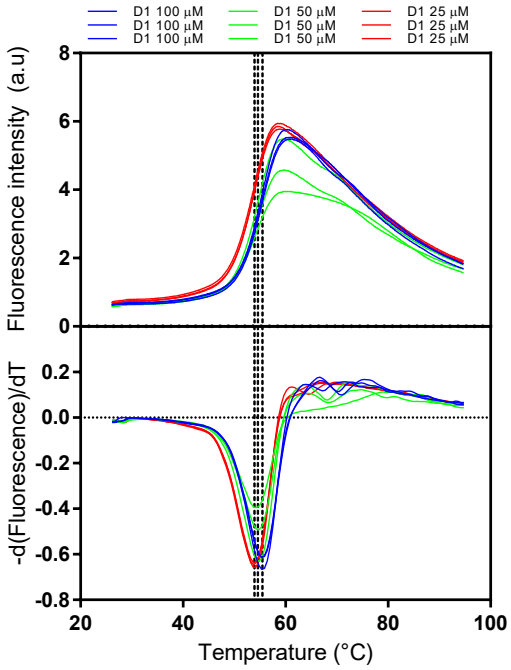
(b) Roche Lightcycler 480 I

Figure 6.2: Thermal melt curves of Cdc20 (161-477) with vehicle control (1% DMSO, black) or without (apo, purple). Each trace represents an individual well. Raw and negative derivative data are represented in top and bottom panels, respectively. (a) Data acquired using the Biorad CFX Connect, with 1.5 µM Cdc20 (161-477), three technical replicates. (b) Data acquired using Roche Lightcycler 480 I using 750 nM Cdc20 (161-477), three technical replicates for Apo, six technical replicates for 1% DMSO. Black dashed lines indicate the T_m as calculated by the average of the minimum from the negative differential plots.

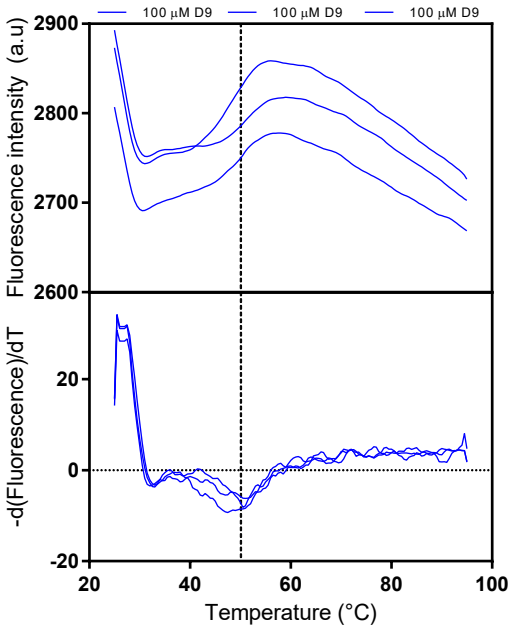
Apcin, which were re-collected as additional positive controls in subsequent experimental runs. Apcin produced a large concentration-dependant change in the T_m of Cdc20 (Figure 6.3a), whereas the Hsl1-derived 'Super D-box' peptide D1 produced a much smaller change in the T_m of Cdc20 (Figure 6.3b), indicating a weaker binding affinity to Cdc20 compared to Apcin. Additionally, the raw data for Cdc20 melt curves with Apcin show a dose-dependant fluorescent intensity baseline of the pre-unfolding transition, which is likely due to the Sypro Orange dye binding to Apcin, which then dissociates with increasing temperature. This is likely due to the low aqueous solubility of the Apcin molecule. In contrast, D1 does not show a similar trend, whereby the pre-unfolding transition baseline is consistent between the concentrations of D1. Peptides D9, A1 and K1 were only screened at the highest concentration of 100 µM initially, due to a lack of Cdc20 material. In all three cases the change in T_m was around 2-3 fold lower than D1 and similar in melting temperature to 1% DMSO, indicative of very weak to no binding. Based on these



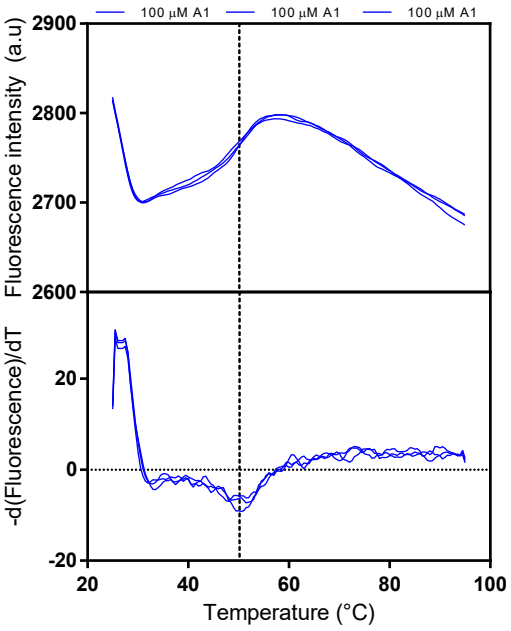
(a) Apcin



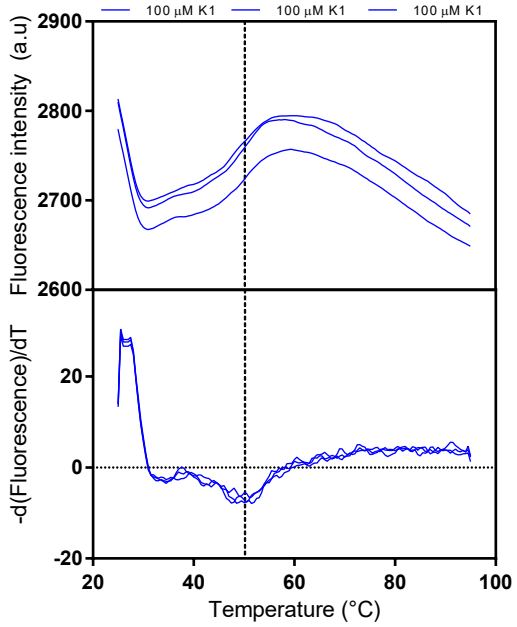
(b) D1: Ac-RAALSDITN-NH_2



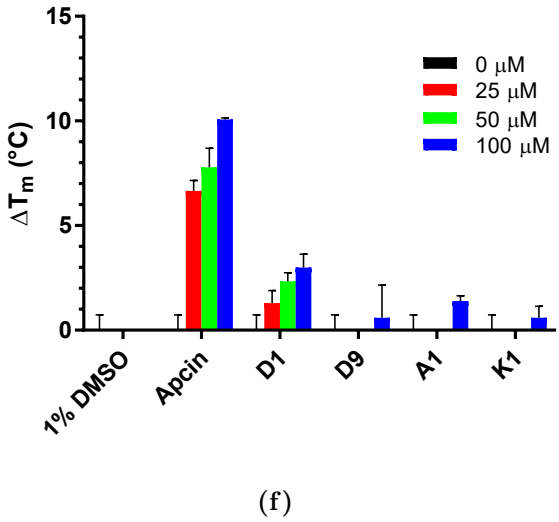
(c) D9: Ac-GRS_5ALSDR$_8$TN-NH_2



(d) A1: Ac-FGLYEE-NH_2



(e) K1: Ac-ELSKENV-NH_2



(f)

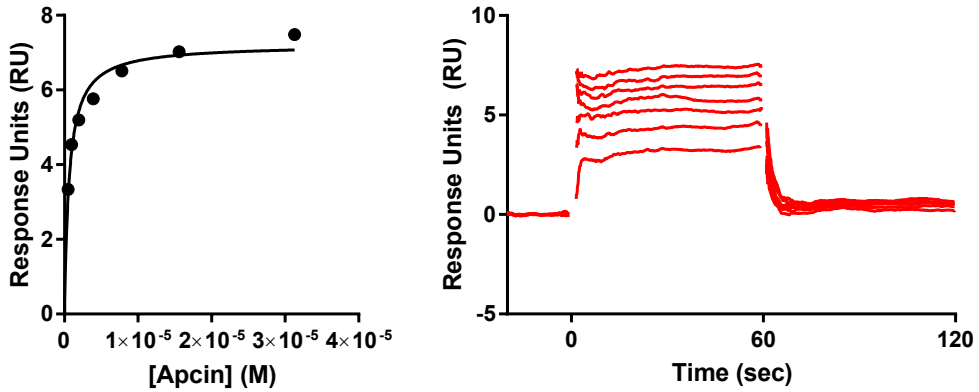
Figure 6.3: Raw and negative derivative data from melt curves of 750 nM Cdc20 (161-477) in the presence of (a) Apcin and (b) D1, at 100 µM (blue), 50 µM (green) and 25 µM (red). Data in (c) D9, (d) A1 and (e) K1 are from melt curves with 1.5 µM Cdc20 (161-477) with 100 µM of the indicated ligand using the Biorad CFX connect instrument. Black dashed lines indicate the average T_m of Cdc20 for the given concentration of ligand. Each trace represents an individual well. (f) Bar charts of the change in melting temperature (ΔT_m) in the presence of different ligands at the indicated concentrations (relative to 1% DMSO). All data are relative to the 1% DMSO, vehicle control data collected from the Biorad CFX connect.

data, and together with the greater wealth of information of D-box degron motifs present in the literature, the ABBA motif and KEN-box peptide inhibitor series were not pursued further.

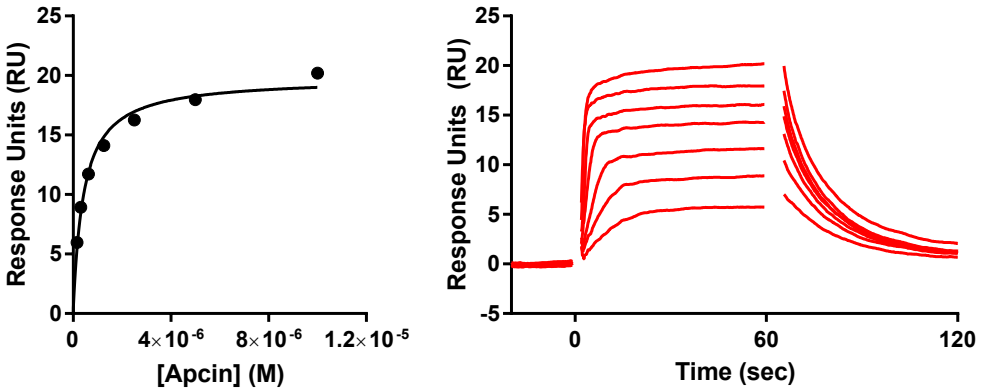
Developing an SPR assay

Since DSF can only provide rough affinity trends, Cdc20-SPR assays were developed alongside. First, Cdc20 was immobilised using Ni-NTA chips since further modification of the protein was not required for this approach. However, given the poor activity of the Cdc20 ligand in this immobilisation strategy, an alternative was explored. The Biotin-Streptavidin interaction is of such high affinity, that subsequent amine coupling, as used in the Ni-NTA immobilisation strategy, is not required. Despite immobilising similar RU of ligand onto both chips (5427 and 4638 RU for the Ni-NTA and SA chips on Fc 4, respectively), the data quality and signal between the two are very different (Figure 6.4 and

Table 6.2). Whereas Ni-NTA captured ligand produced only poor quality sensorgrams, the SA-captured ligand exhibited clean association and dissociation phases.



(a) Ni-NTA capture: Apcin Fc4-1



(b) SA capture: Apcin Fc4-1

Figure 6.4: Difference in raw data quality between ligand immobilisation using (a) Ni-NTA capture or (b) SA capture. Steady state affinity fits and sensorgrams are shown for Fc 4-1 for the small molecule Apcin. Results from data fitting are listed in Table 6.2.

Furthermore, an approximate 3-fold increase of signal is observed using the SA chip compared to the Ni-NTA chip. Although dissociation constants are somewhat similar, the fit from SA immobilised protein is more reliable due to the increase raw data quality. The amount of functional Cdc20 ligand on the Ni-NTA chip is very low, indicating that the ligand is either binding incompetent while remaining folded, e.g. by being attached to the surface in an orientation that occludes the binding site, or that it has lost its native fold during the immobilisation procedure. For example, due to multiple freeze-thaw cycles

Table 6.2: Results from data fitting of steady-state affinity measurements of Apcin using a Ni-NTA or SA chip.

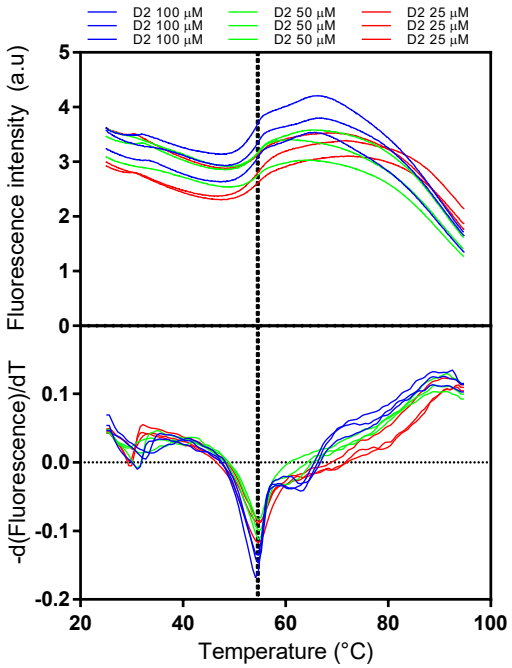
Biosensor chip	Ni-NTA	SA
RU_{max} (RU)	7.2	19.8 ± 0.6
K_d (nM)	634 ± 96	418 ± 53
% $Ligand_{functional}$	11.9	34.4

and sample preparation, aggregates or misfolded monomers could accumulate that may still bind the chip but are unable to interact with the analyte. However, if around 60% of the protein was not folded correctly, DSF melt curves would exhibit an altered transition profile. Since abnormal thermal shifts were not observed, it is more likely that the protein remains largely folded correctly, but can simply not bind the analyte for various reasons. All dissociation constants quoted in the following sections were generated using biotinylated Cdc20 ligand immobilised onto the SA chip.

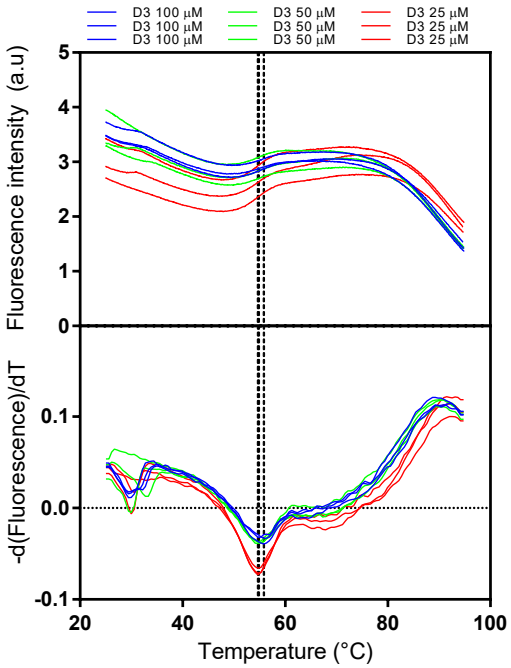
6.3.2 Exploring position 7 of the D-box motif

The amino acids found at position 7 of the D-box peptide motif tend to be aliphatic (valine, isoleucine and leucine, Figure 5.4) [165]. Peptides D3 (Ac-RAPLGDVSN-NH_2), D4 (Ac-RAPLGDISN-NH_2) and D5 (Ac-RAPLGDLSN-NH_2) were designed and synthesised incorporating point mutations of valine, isoleucine and leucine at position 7, respectively. In addition, the consensus D-box motif, peptide D2 (Ac-RLPLGDVSN-NH_2) was synthesised and tested as well. Peptides were initially screened for binding by DSF (Figure 6.5).

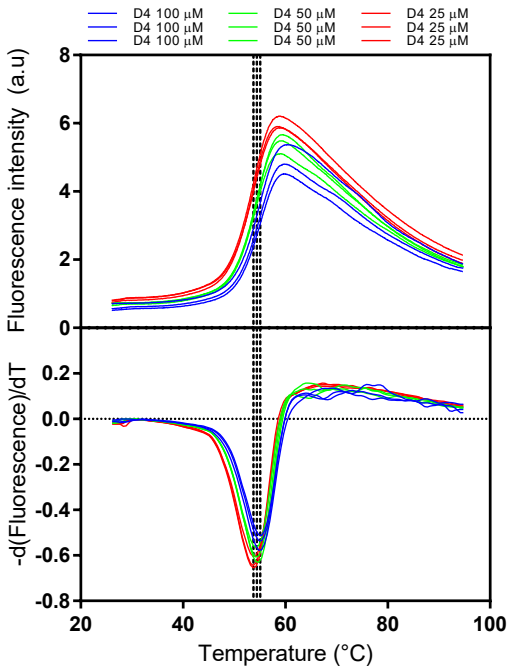
DSF results for all four peptides are shown in Figure 6.5, with Apcin included as a positive control molecule that can also serve as an internal standard between different sets of experiments. However, the T_m shift of Cdc20 with 100 µM Apcin appeared to be higher in the first run (+10.08 ± 0.05 °C, Figure 6.3f) than in the second run (+8.08 ± 0.08°C, Figure 6.5). Indeed, subsequent DSF experiments show the second result to be more representative of the average, and the higher observed T_m initially could be a result of poor sample dilution. Thermal shifts due to D1 appear more consistent than for Apcin between the first and second experiments (2.99 ± 0.37°C and 2.89 ± 0.06°C, respectively). Surprisingly, little to no evidence of binding is observed for peptide D2, the consensus sequence. Peptides D3, D4 and D5 show dose-dependent T_m shifts, which are all smaller than that of D1 at the same concentrations. As was shown for Apcin and D1 in Figures 6.3a and 6.3b, the raw data for D4 is shown for a subsequent experiment where



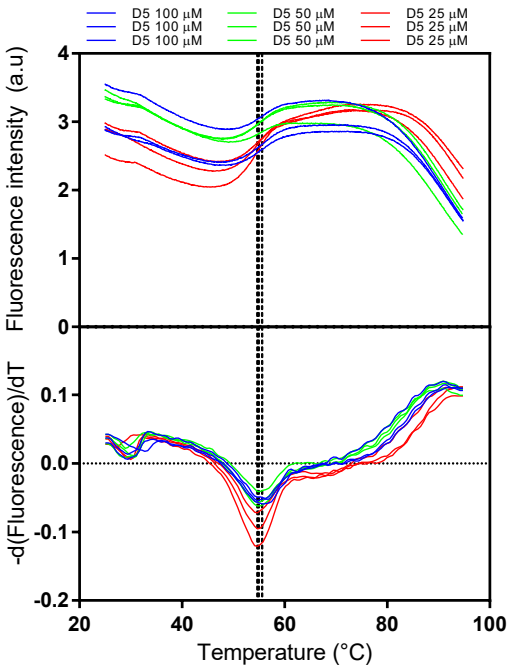
(a) D2: Ac-RLPLGDVSN-NH_2



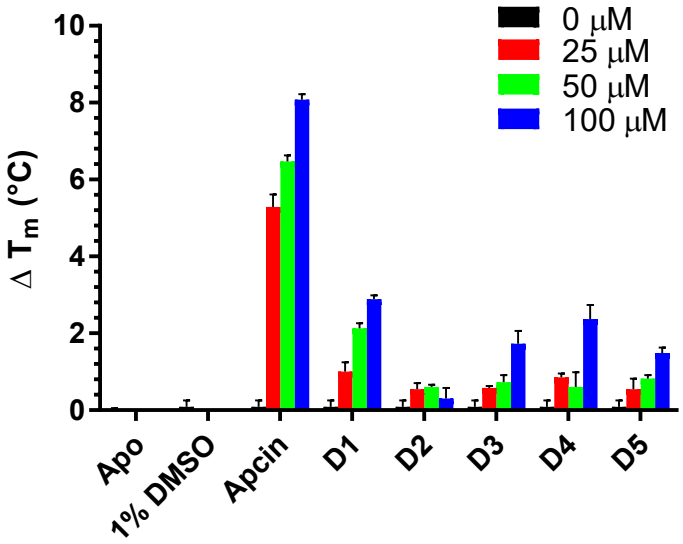
(b) D3: Ac-RAPLGDVSN-NH_2



(c) D4: Ac-RAPLGDISN-NH_2



(d) D5: Ac-RAPLGDLSN-NH_2

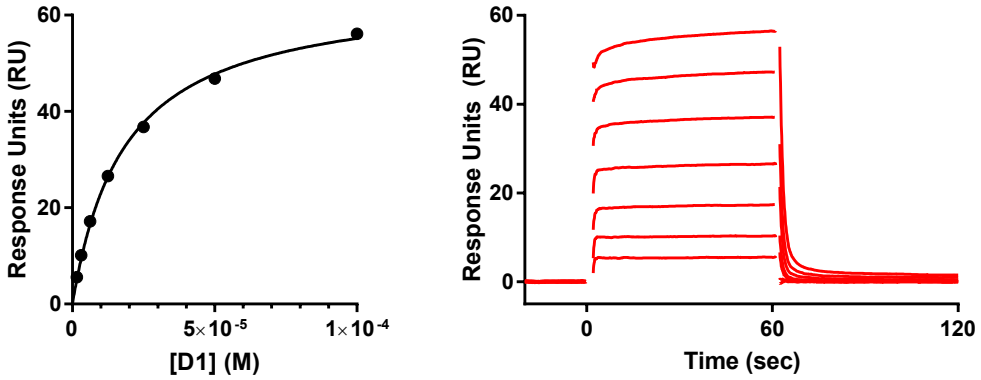


(e)

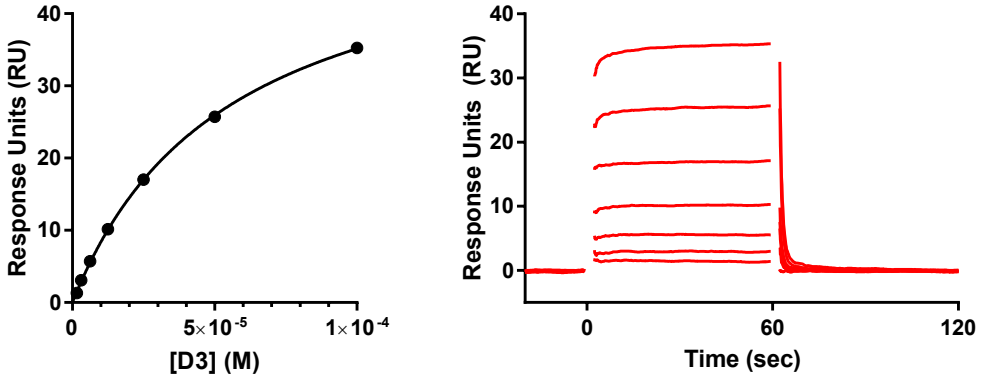
Figure 6.5: Raw and negative derivative data from melt curves of 750 nM Cdc20 (161-477) in the presence of (a) D2 and (b) D3, (c) D4 and (d) D5 at 100 µM (blue), 50 µM (green) and 25 µM (red). Black dashed lines indicate the average T_m of Cdc20 for the given concentration of ligand. Each trace represents an individual well. Raw data for D4 is representative from a subsequent experiment, where the overall quality was improved. (e) Bar charts of the change in melting temperature (ΔT_m) in the presence of different ligands at the indicated concentrations (relative to 1% DMSO). All data are relative to the 1% DMSO, vehicle control data collected on the Roche Lightcycler 480 I.

generated better quality data, whereas peptides D2, D3 and D5 were not subsequently repeated due to lack of sufficient purified Cdc20. Since DSF data did not provide suitable resolution for the relatively small changes in T_m of Cdc20, these experiments were not entirely sufficient to determine which residue at position 7 was optimal.

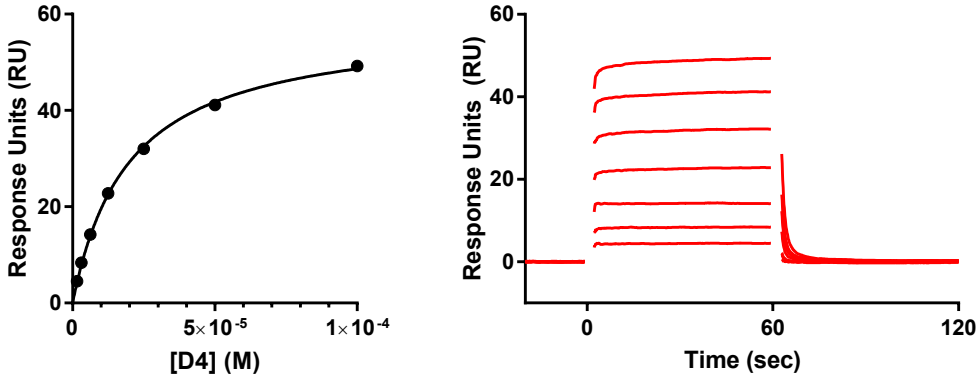
All peptides except D2, for which there was no clear evidence of binding (Figure 6.5), were further analysed by SPR. The data presented in Figure 6.6 and Table 6.3 show a clear preference for longer aliphatic side chains: isoleucine and leucine gave higher affinities than the shorter side chain of valine. In addition, isoleucine is better than leucine and it was therefore incorporated into all subsequently synthesised D-box peptides at position 7. Peptides D1 and D4 have very similar dissociation constants, and this similarity is also apparent in the DSF data (Figure 6.5).



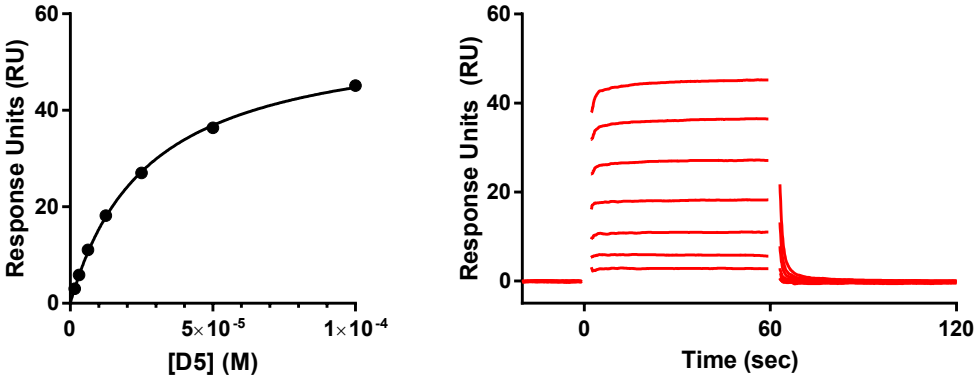
(a) Fc4-1, D1: Ac-GRAALSDITN-NH_2



(b) Fc4-1, D3: Ac-RAPLGDVSN-NH_2



(c) Fc4-1, D4: Ac-RAPLGDISN-NH_2



(d) Fc4-1, D5: Ac-RAPLGDLSN-NH_2

Figure 6.6: Representative data from Fc 4-1 of SPR steady state affinity fit plots (left panels) and sensorgrams (right panels) of peptides D1, D3, D4 and D5. Results from data fitting are listed in Table 6.3.

Table 6.3: SPR analysis of Cdc20 binding to D3, D4 and D5. K_d values are reported as the average of all three reference-subtracted flow cells, with the standard error of the mean. RU_{max} and % $Ligand_{functional}$ are reported for Fc4-1 only.

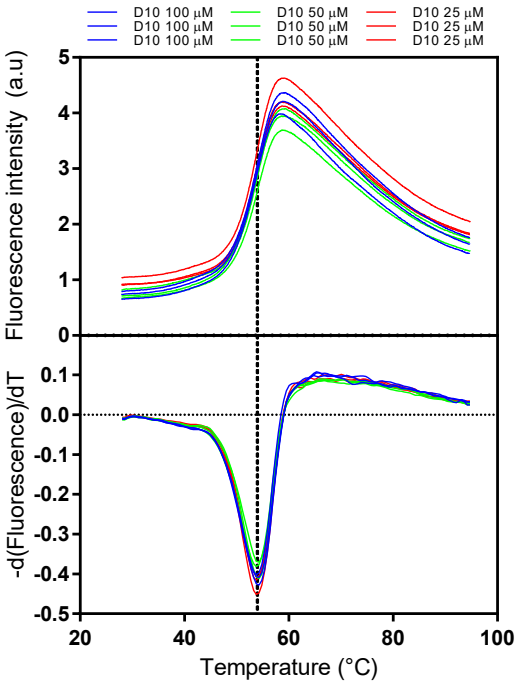
Peptide	D1	D3	D4	D5
RU_{max} (RU)	65.29	54.5	58.25	56.9 ± 0.6
K_d (μM)	18.59 ± 0.22	54.37 ± 0.74	19.58 ± 0.22	27.12 ± 1.04
% $Ligand_{functional}$	47.0	42.8	45.1	44.0

6.3.3 Exploring position 3 of the D-box motif

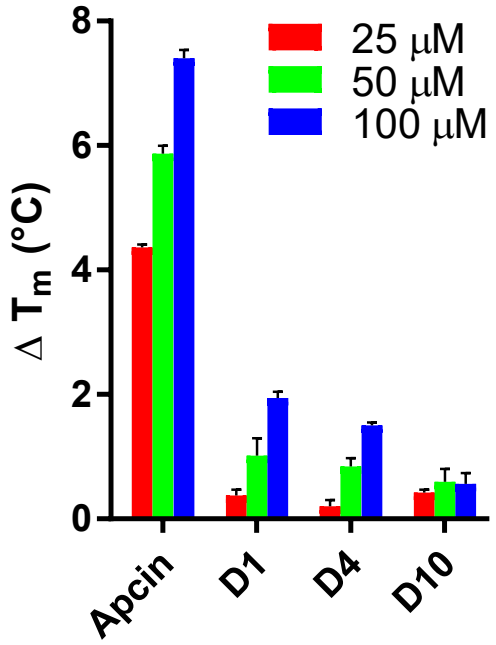
The consensus sequence of D-box peptides has a strong preference for either proline or alanine at position 3, with proline being observed at a higher frequency. To investigate which of the two amino acids at position 3 were optimal for binding, a derivative of peptide D4 (Ac-RAPLGDISN-NH_2) was synthesised to introduce alanine in its place, resulting in peptide D10 (Ac-RAALGDISN-NH_2). Indeed, the inclusion of alanine at position 3 of the D-box peptide is not favourable as evident from the reduced thermal stabilisation of Cdc20 compared to peptide D4 (Figures 6.7a and 6.7b) and a 3.5-fold weaker binding affinity as measured by SPR (69.56 ± 3.39 μM Figure 6.7c).

6.3.4 Exploring positions 5 and 8 of the D-box motif

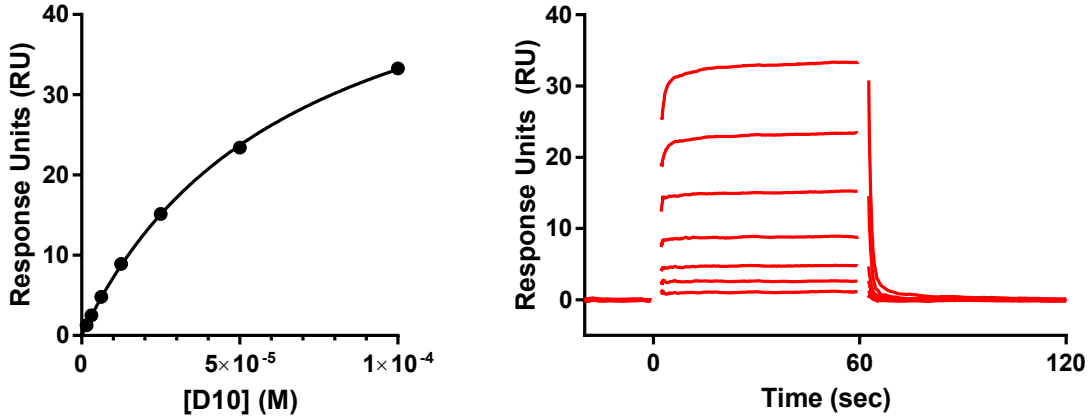
The remaining differences between peptides D1 (Ac-GRAALSDITN-NH_2) and D4 (Ac-RAPLGDISN-NH_2), excluding the solvent-exposed N-terminal glycine residue of D1, are



(a) D10: Ac-RAALGDISN-NH_2



(b)

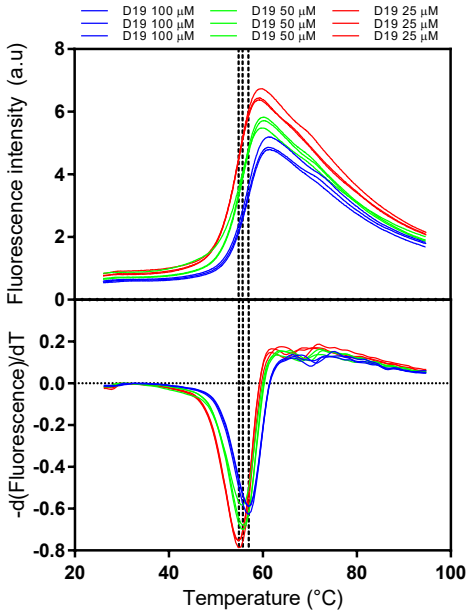


(c) Fc4-1, D10: Ac-RAALGDISN-NH_2

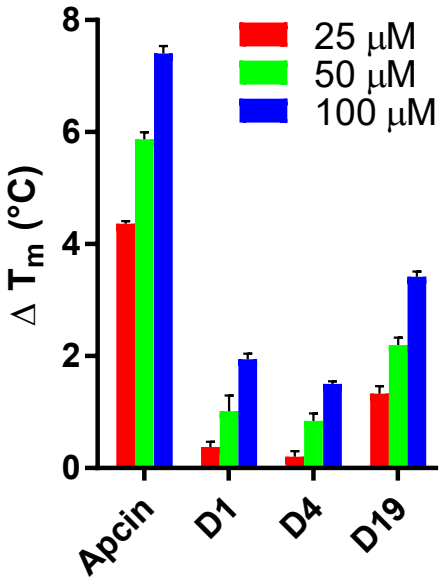
Figure 6.7: (a) Raw and negative derivative data from melt curves of 750 nM Cdc20 (161-477) in the presence of D10 at 100 µM (blue), 50 µM (green) and 25 µM (red). Black dashed lines indicate the average T_m of Cdc20 for the given concentration of ligand. Each trace represents an individual well. (b) DSF analysis of D10 compared to D1 and D4 and Apcin binding to Cdc20. (c) Representative SPR data of D10 binding to Cdc20 on Fc4-1. Steady state affinity (left) and sensorgram (right).

positions 5 and 8 (highlighted in red). Due to a lack of time to explore both positions independently, peptide D19 (Ac-RAPLSDITN-NH_2) was synthesised, incorporating both residues from D1 into the D4 sequence. Only 25 µM of D19 produces a similar response

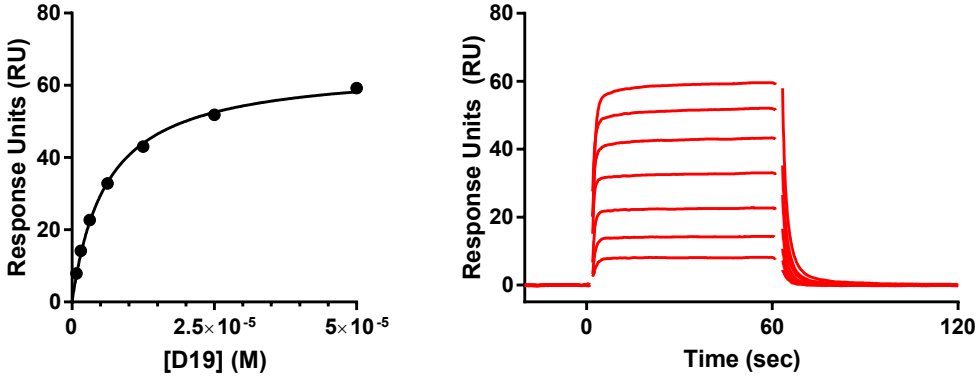
in DSF compared to 100 µM of D1 or D4 alone (Figure 6.8b), and D19 can be observed to bind with a dissociation constant of 5.87 ± 0.11 µM by SPR (Figure 6.8c). Despite deriving an optimal sequence comprising natural amino acids, it was evident that the affinity of a natural peptide was not comparable to Apcin (Table 6.2).



(a) D19: Ac-RAPLSDITN-NH_2



(b)

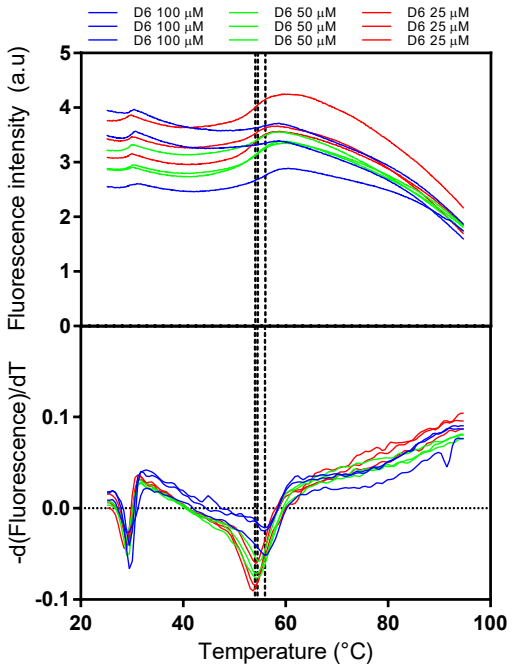


(c) Fc4-1, D19: Ac-RAPLSDITN-NH_2

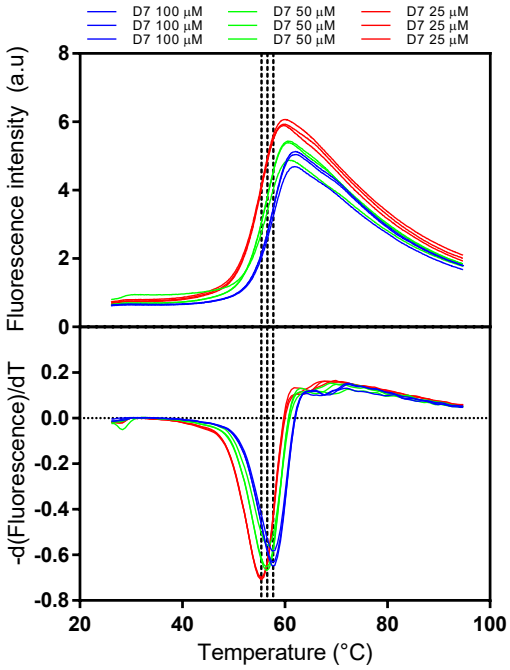
Figure 6.8: (a) Raw and negative derivative data from melt curves of 750 nM Cdc20 (161-477) in the presence of D19 at 100 µM (blue), 50 µM (green) and 25 µM (red). Black dashed lines indicate the average T_m of Cdc20 for the given concentration of ligand. Each trace represents an individual well. (b) DSF analysis of D19 compared to D1 and D4 and Apcin binding to Cdc20. (c) Representative SPR data of D19 binding to Cdc20 on Fc4-1. Steady state affinity (top) and sensorgram (bottom).

6.3.5 Incorporating non-natural amino acids at position 4

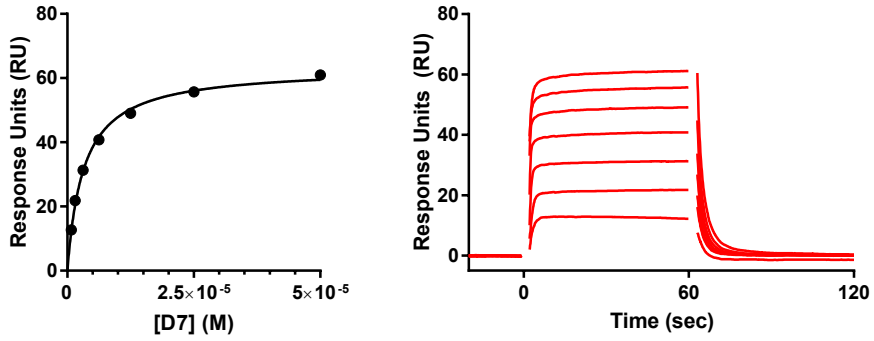
Initially, (*S*)-2-amino-3-cyclobutylpropanoic acid (cBuAla), which had previously been used in the context of the p53 peptide [105], and (*S*)-2-amino-4,4-dimethyl-pentanoic acid (C_3) were explored as functional derivatives of the leucine side chain (for structures see Figure 5.8). It was hypothesized that burial of a slightly bulkier side chain into the D-box pocket may increase the affinity. Both amino acids were incorporated into the scaffold of D4, resulting in the synthesis of peptides D6 (Ac-RAPcBuAlaGDISN-NH_2) and D7 (Ac-RAPC_3GDISN-NH_2). Although substitution of cBuAla did not yield an improved ligand, peptide D7 showed significant improvement over peptide D4, with a change in melting temperature of 4.13 ± 0.03°C versus 1.51 ± 0.03°C, respectively (Figure 6.9).



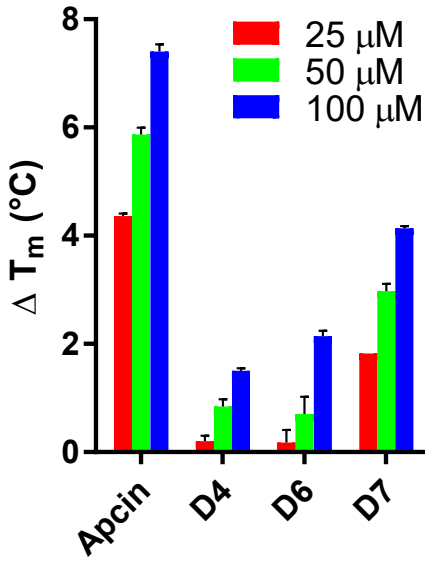
(a) D6: Ac-RAPcBuAlaGDISN-NH_2



(b) D7: Ac-RAPC_3GDISN-NH_2



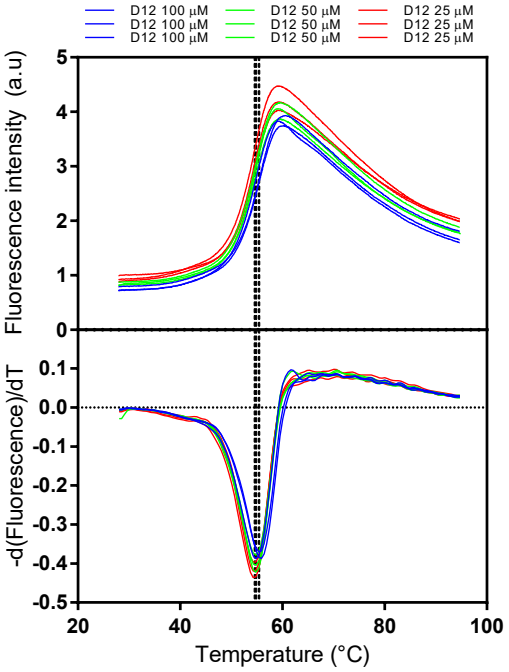
(c) Fc4-1, D7: Ac-RAPC_3GDISN-NH_2



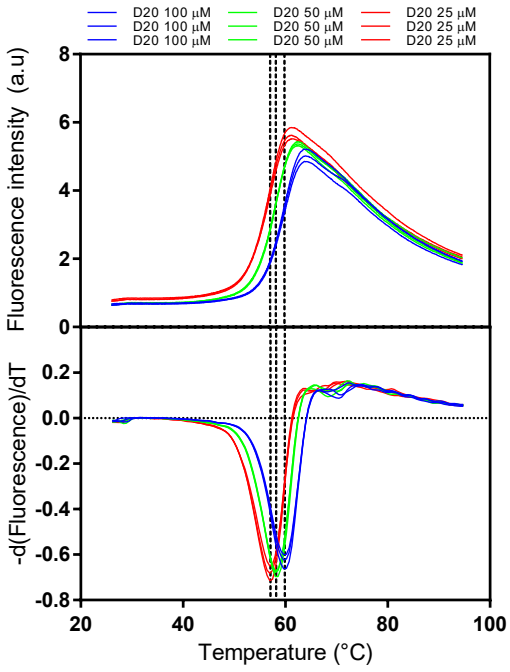
(d)

Figure 6.9: Raw and negative derivative data from melt curves of 750 nM Cdc20 (161-477) in the presence of (a) D6 and (b) D7 at 100 µM (blue), 50 µM (green) and 25 µM (red). Black dashed lines indicate the average T_m of Cdc20 for the given concentration of ligand. Each trace represents an individual well. (b) Representative SPR data of D7 binding to Cdc20 on Fc4-1. Steady state affinity (left) and sensorgram (right). (c) DSF analysis of D6 and D7 compared to Apcin and the parental peptide sequence, D4, binding to Cdc20.

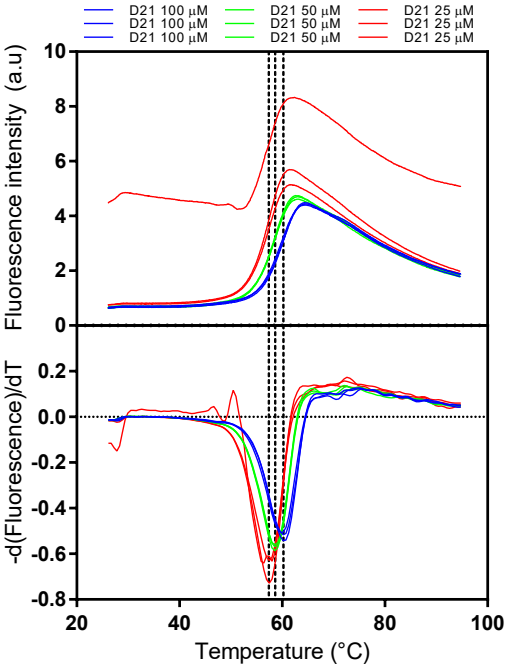
These results indicated that the higher affinity could be driven by the leucine-binding pocket of Cdc20. The C_3 amino acid was substituted into peptides D10 and D19 (creating D12 and D20, respectively) to see if it could improve the binding of both the lower- and higher-affinity peptide scaffolds to Cdc20. In all three C_3-containing derivative peptides, increases in Cdc20 thermal stability and binding affinity were observed compared to their leucine-containing parental sequences (Figures 6.9 and 6.10 and Table 6.4). Surprisingly, the 6-fold increase in affinity upon incorporating (*S*)-2-amino-4,4-dimethylpentanoic acid appeared to be independent of the parental sequence. Attempts to make the peptide more similar to Apcin by incorporation of a halogenated amino acid were also pursued. As mentioned in Chapter 5, Fmoc-(*S*)-2-amino-4,4,4-trichlorobutanoic could not be produced by a commercial vendor. The closest halogenated alternative, Fmoc-(*S*)-2-amino-4,4,4-trifluorobutanoic acid, is commercially available and was used to synthesise peptide D21 (Ac-RAPF_3SDITN-NH_2), which exhibits the tightest binding to Cdc20 of all peptides by producing an 11-fold increase in affinity (Figure 6.10 and Table 6.4).



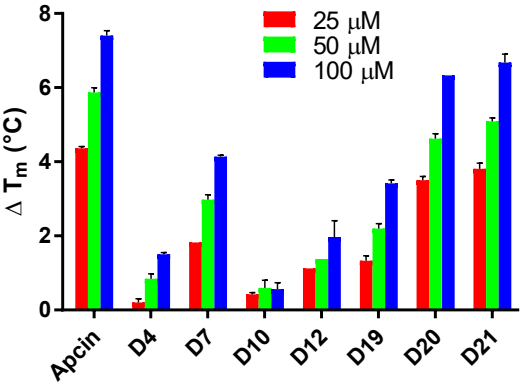
(a) D12: Ac-RAAC$_3$GDISN-NH$_2$



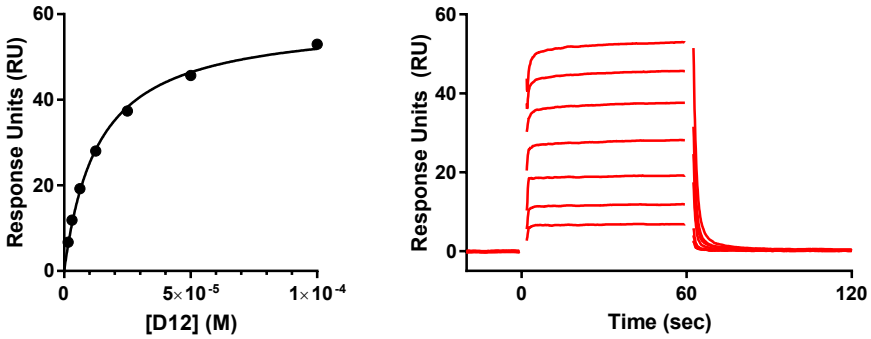
(b) D20: Ac-RAPC$_3$SDITN-NH$_2$



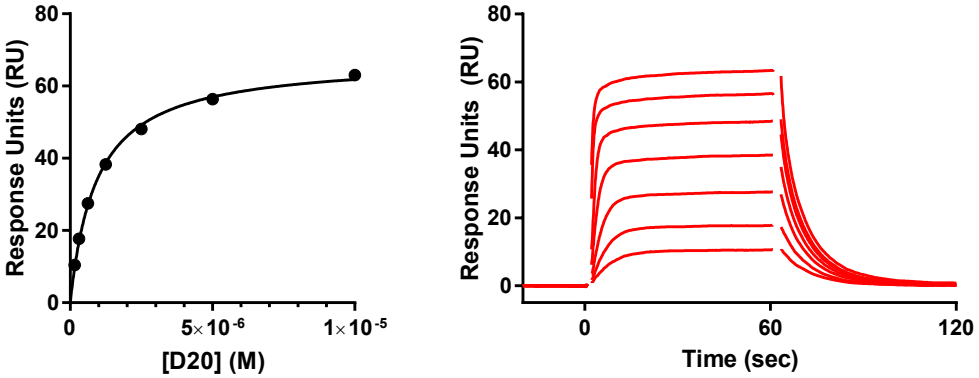
(c) D21: Ac-RAPF$_3$SDITN-NH$_2$



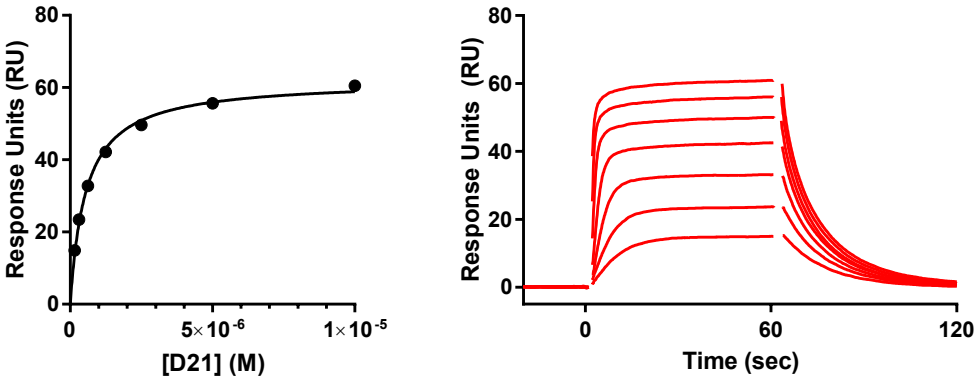
(d)



(e) Fc4-1, D12: Ac-RAAC_3GDISN-NH_2



(f) Fc4-1, D20: Ac-RAPC_3SDITN-NH_2



(g) Fc4-1, D21: Ac-RAPF_3SDITN-NH_2

Figure 6.10: Biophysical characterisation of peptide derivatives containing a substitution of (*S*)-2-amino-4,4-dimethylpentanoic acid (C_3) and (*S*)-2-amino-4,4,4-trifluorobutanoic acid (F_3), for leucine at position 4. Raw and negative derivative data from melt curves of 750 nM Cdc20 (161-477) in the presence of (a) D12, (b) D20 and (c) D21 at 100 µM (blue), 50 µM (green) and 25 µM (red). Black dashed lines indicate the average T_m of Cdc20 for the given concentration of ligand. Each trace represents an individual well. (d) DSF analysis of D12, D20 and D21 compared to Apcin and their respective parental peptide sequences, binding to Cdc20. SPR steady state affinity fits (Left) and sensorgrams (Right) for peptides (e) D12, (f) D20 and (g) D21.

Table 6.4: Fitting outputs of steady-state affinity measurements Fc 4-1 of D4, D7, D10, D12, D19, D20 and D21. K_d is the average of all three reference-subtracted flow cells with the standard error of the fits.

Peptide	RU_{max} (RU)	K_d	% $Ligand_{functional}$	Fold K_d increase
D4	58.25	19.58 ± 0.22 μM	45.1	-
D7	63.31	3.10 ± 0.10 μM	48.3	6.3
D10	55.22	69.56 ± 3.39 μM	43.9	-
D12	58.88	13.30 ± 0.10 μM	46.1	5.2
D19	65.24	5.87 ± 0.11 μM	48.3	-
D20	67.55	902 ± 14 nM	49.4	6.5
D21	62.07	524 ± 14 nM	44.9	11.2

Most notably, the thermal stabilisation of Cdc20 by Apcin, is now almost achieved by both the D20 and D21 peptides (Figure 6.10d). For the first time, we can directly make a qualititative comparison on the correlation of thermal stabilisation by a Cdc20 ligand with affinities derived from SPR steady-state equilibrium fits. Though it could also be inferred similarly with the weaker ligands, such as D1, D3, D4 and D5, the resolution and clarity of this was more difficult to observe, hence the defined the necessity for the SPR experiments. Indeed, from the DSF data there is a clear discrepancy for order of thermal stabilisation with Apcin > D21 > D20. This trend observed by DSF also follows for their calculated affinities of 419 ± 53 nM, 524 ± 14 nM and 902 ± 14 nM, respectively. This observation, particularly in the higher affinity regime demonstrates that DSF can be a high-throughput and reliable method for screening ligands targeting Cdc20.

6.3.6 Cellular Thermal Shift Assays

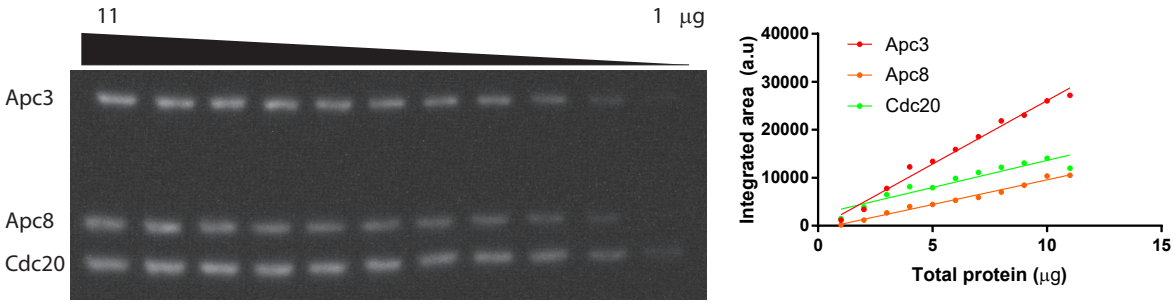
Cellular Thermal Shift Assays (CETSA), a relatively new technique similar to DSF, relies on the fact that ligand binding will influence the melting temperature (T_m) of a protein. CETSA has the benefit of working in a more physiological context, such as mammalian cells or even tissues [244], instead of being reliant on recombinant protein in aqueous buffers. Here, the stabilisation of both endogenous and exogenous target proteins was explored in the background of HEK 293T lysates. Working with endogenous protein has a major biological advantage over transiently transfected (exogenous) protein. First, protein levels are reflective of the natural system instead of being overexpressed. Second, it is not clear that exogenous protein will have the same functional capability as endogenous protein. However, detection of endogenous protein levels and changes therein can be difficult, whereas exogenous proteins can be modified to include small tags for various detection methods.

CETSA of endogenous proteins

Since changes in melting temperature observed in DSF experiments were sufficiently large, it was hypothesised that a ligand would also stabilise endogenous Cdc20 in the cell as measured by CETSA. Additionally, it was thought that a similar approach could be used to investigate ligand binding to Apc8 and Apc3, for which it was not feasible to establish assays using purified protein. A dilution series of endogenous HEK 293T lysate was analysed to probe the linearity of the densitometric analysis and was which was found to be approximately linear for Apc3, Apc8 and Cdc20 (Figure 6.11). Sackton *et al.* [195] performed an isothermal melt using a titration of ligand to induce increased stabilisation of Cdc20. However, attempts to recapitulate the isothermal CETSA data by western blot were unsuccessful (data not shown) and therefore a thermal melt curve approach was used instead. The addition of 100 µM Apcin was sufficient to stabilise Cdc20, whereas 100 µM TAME induced stabilisation of Apc3 but not Apc8 (Figure 6.12 and Table 6.5).

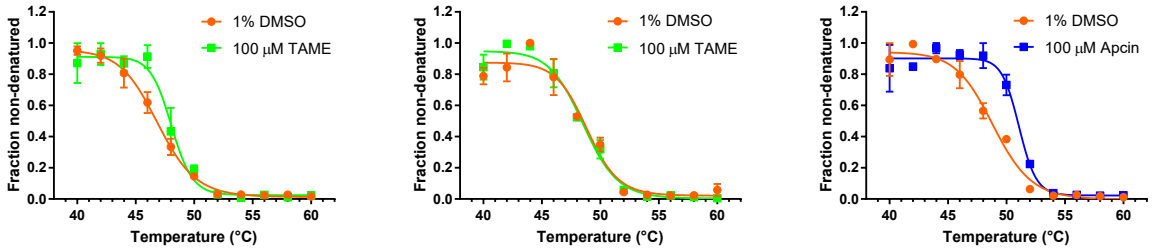
Table 6.5: Melting temperatures of Apc3, Apc8 and Cdc20 from endogenous HEK 293T lysates, with small molecule compounds.

Protein	Melting temperature (°C)		
	1% DMSO	TAME	Apcin
Apc3	46.88 ± 0.27	48.07 ± 0.26	-
Apc8	48.88 ± 0.37	48.57 ± 0.29	-
Cdc20	48.88 ± 0.29	-	51.09 ± 0.21



(a) **(b)**

Figure 6.11: (a) Dilution of HEK 293T lysate from 11 to 1 µg of protein loaded onto each lane. (b) Densitometric analysis of the HEK 293T dilution series.



(a) Apc3 **(b)** Apc8 **(c)** Cdc20

Figure 6.12: Thermal melt curves of proteins from endogenous HEK 293T lysates. (a) Apc3 with 100 µM TAME (green) or vehicle control (orange) (b) Apc8 with 100 µM TAME (green) or vehicle control (orange). (c) Cdc20 with 100 µM Apcin (blue) or vehicle control (orange). Data represent the average of biological duplicates.

CETSA of exogenous Cdc20

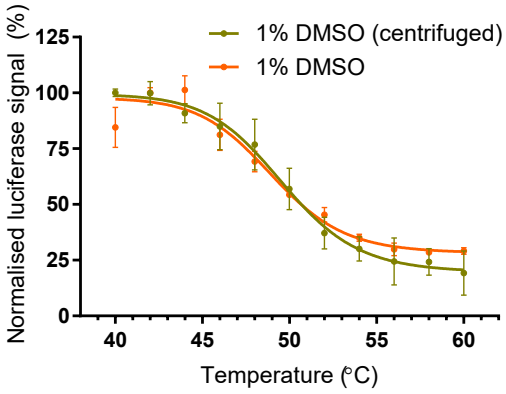
Quantitative western blotting is particularly low in throughput and, therefore, a split-luciferase detection method was implemented [237]. This approach utilises the LgBiT protein and the HiBiT tag, a high-affinity complementing peptide (VSGWRLFKKIS) that is able to reconstitute the NanoLuc luciferase (Promega). The HiBiT-tag (with additional GS-linkers appended to both ends of the tag) was added to full-length Cdc20 (1-499) at either the N- or C- terminus. In a preliminary experiment, the signal for the Cdc20 (1-499)-HiBiT (C-terminal) construct was found to be significantly better than the N-terminally tagged protein and therefore all future experiments were conducted using the C-terminally tagged variant. The 1% DMSO control sample was used to examine whether the removal precipitated protein after heating by centrifugation had a significant impact on data quality (Figure 6.13a). Cdc20 in the centrifuged sample showed a T_m of 49.56 ± 0.70°C, whereas Cdc20 in the non-clarified sample showed a T_m of 48.98 ± 0.43°C. Since these values are the same within error, the centrifugation step could be omitted to speed up sample processing, highlighting one major advantage of the split-luciferase CETSA

method compared to western blot detection. Furthermore, melting temperatures of exogenous protein were similar to the T_m of endogenous Cdc20 and this provided confidence that the protein was not behaving significantly differently when overexpressed (Table 6.5). Next, Apcin and the top four D-box peptides (D7, D19, D20 and D21) were tested to see if they stabilised (i.e. bound to) the exogenous Cdc20 protein (Figure 6.13b-f) and Table 6.6).

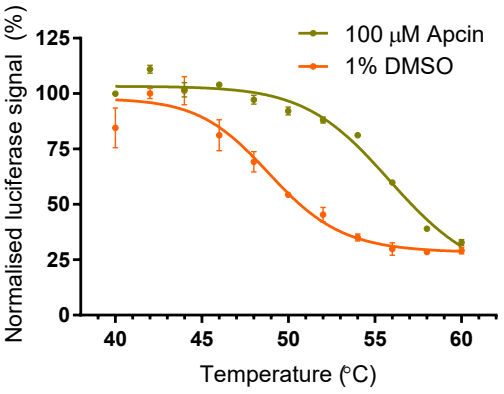
With all novel D-box peptide ligands, we observed a significant stabilisation of Cdc20 relative to the vehicle control. However, the weakest ligand, D19, produces only a marginal positive shift of 0.9°C, highlighting that ligands in the low-to-mid micromolar affinity range are those at the detection limit of the developed system. Additionally, despite D21 and Apcin having very similar affinities as measured by SPR and DSF, there is a larger discrepancy in the relative melting temperature shift of both compounds in this experiment. Sackton and co-workers demonstrated Apcin's specificity to binding Cdc20 in comparison to other WD40 domain proteins, including the Cdc20 homologue protein, Cdh1 [195]. Due to a lack of time, experiments to this extent were not performed and it is indeed very likely that these D-box ligands will also have similar affinities to Cdh1.

Table 6.6: Melting temperatures of Cdc20 (1-499)-HiBiT in transfected HEK 293T lysates, with D-box ligands.

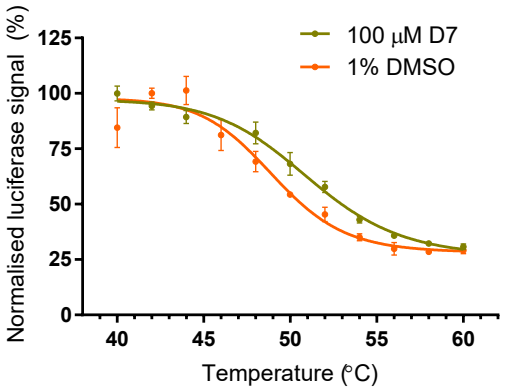
Ligand	Melting temperature (°C)
1% DMSO	48.98 ± 0.43
100 μM Apcin	56.06 ± 0.70
100 μM D7	50.94 ± 0.40
100 μM D19	50.88 ± 0.31
100 μM D20	53.06 ± 0.44
100 μM D21	53.42 ± 0.37



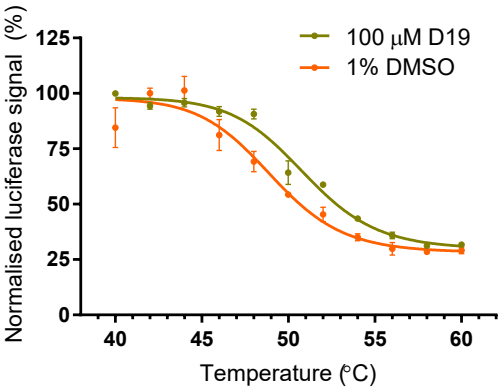
(a) Vehicle control



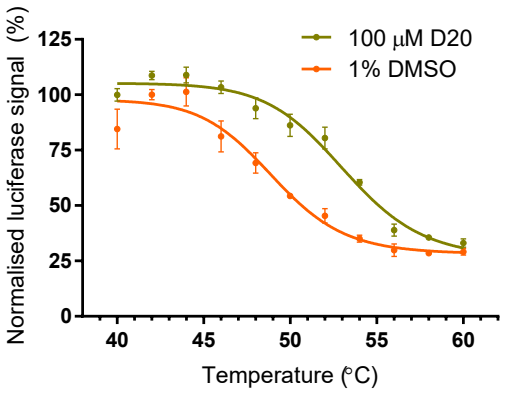
(b) Apcin



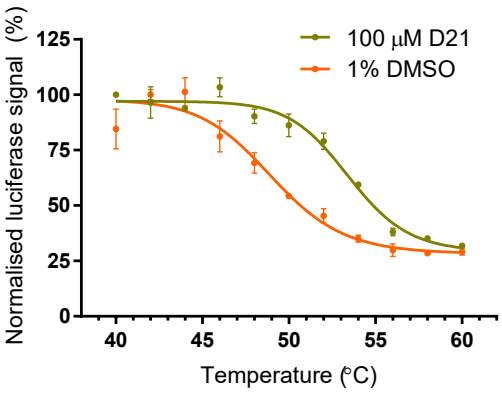
(c) D7: Ac-RAPC_3GDISN-NH_2



(d) D19 Ac-RAPLSDITN-NH_2



(e) D20: Ac-RAPC_3SDITN-NH_2



(f) D21: Ac-RAPF_3GDISN-NH_2

Figure 6.13: Thermal melt curves of exogenous Cdc20 (1-499)-HiBiT in the presence of vehicle control (1% DMSO), Apcin, D7, D19, D20 or D21. (a) Impact of centrifugation as determined by comparing centrifuged (lime-green) and non-centrifuged (orange) samples of 1% DMSO control. Ligand-induced stabilisation of Cdc20 using (b) 100 µM Apcin, (c) 100 µM D7, (d) 100 µM D19, (e) 100 µM D20 or (f) 100 µM D21. All samples are overlaid with the 1% DMSO control . Data represent the average of triplicate experiments.

CETSA of exogenous Apc8

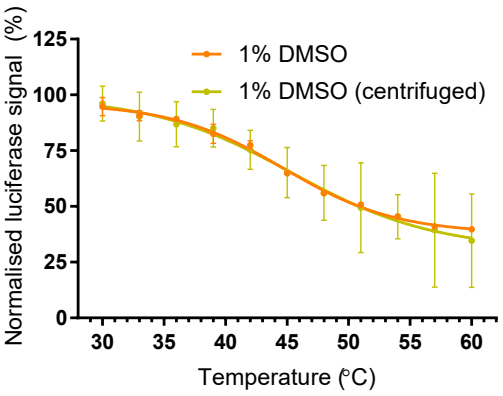
Although we had demonstrated that we could observe binding of TAME to endogenous Apc3 by the western blotting methodology (Figure 6.12), it still suffers from the same low-throughput and low sensitivity as was seen with CETSA experiments with Cdc20. Due to this fact, developing the same high-throughput split-luciferase assay format was desirable for both Apc3 and Apc8. However, for reasons that were not clear, very poor expression of the HiBiT-Apc3 (1-824) and Apc3 (1-824) constructs were observed, which were only 7-fold greater than the background signal observed of untransfected lysates. It is possible that the introduction of the HiBiT tag at either terminus of Apc3 hinders its' incorporation into the APC/C complex and thus may be very unstable as a result, as indicated from personal communication with our collaborator, Prof. David Barford, an expert in the structural biology of the APC/C. Unperturbed by the poor signal from the HiBiT-tagged Apc3 constructs, we explored HiBiT-tagged Apc8 constructs. In the preliminary experiments, the C-terminally tagged construct showed approximately twice the raw luminescence values compared to the N-terminally tagged construct. Here, we observed greater than 100-fold signal change over the background of untransfected cell lysates indicating a more stable construct that is assumed to be incorporated in the APC/C complex. First, we demonstrated that the Apc8 (1-591)-HiBiT construct would also behave similarly as we could see with the Cdc20 constructs, whereby it was possible to remove the centrifugation step without significantly affecting the melting temperature of Apc8 (Figure 6.14a). Having established that the T_m of the non-centrifuged and centrifuged samples were 45.3 $\pm$ 0.5 and 44.3 $\pm$ 0.3, (Table 6.7), respectively as an average from experimental triplicates. While the difference here is greater than the standard deviation from the experimental triplicates, curve fitting of the centrifuged samples is more erroneous given the larger error between technical replicates of each temperature point, where it is believed that the centrifugation step in this process leads to increased heterogeneity across the technical replicates in each assay plate.

The focus then turned to testing the TAME small molecule alongside the C-box peptides to observe if they would stabilise Apc8. Observing the TAME small molecule in this split-luciferase assay format (Figure 6.14b) is in agreement with what was observed in the initial western blotting experiments looking at TAME against endogenous Apc8 (Figure 6.12b). In both experiments, no thermal stabilisation of Apc8 was observed and this provided some confidence that TAME was indeed a poor inhibitory ligand as has been described in the literature, when used on it's own [194, 195]. Observing both the linear, C1, and Alanine-mutant knockout peptide, C5, we similarly observe little to no change in the melting temperature of the reporter construct. While we expected the neg-

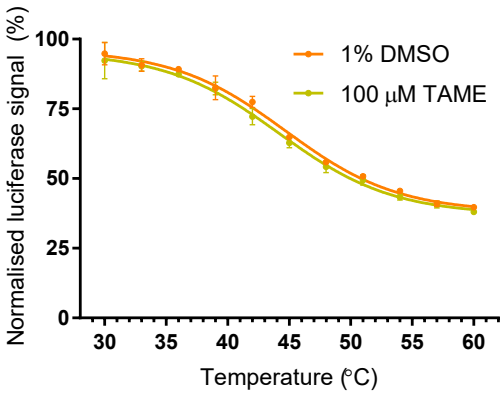
ative control peptide, C5, to do exactly this, the same was not expected for the linear peptide. Since this assay format is the only way that has been explored to provide a semi-biophysical interpretation of ligand binding to the target protein, it may well be possible that the affinity is too low to be detectable. Indeed, as was the case with Cdc20 and D-box ligands, the mid-single digit micromolar range appeared to be the cut-off of detectable thermal stabilisation in those CETSA experiments. However, in stark contrast to the other ligands, both the early and later isomers of the hydrocarbon-stapled peptide ligand, C6e and C6l, produced a significant destabilising effect to exogenous Apc8. Although rarely observed in drug discovery high-throughput screening campaigns, both ligands that produce thermal stabilisation or destabilisation are of interest. Destabilising ligands may offer mechanisms of allosteric inhibition by shifting the global protein structure and thereby reducing the free energy required to unfold the protein. However, in this specific case, since we are working in a non-purified system, it is perhaps more likely that we have a element of competition for the C-box binding site. In the cell lysate, either Cdc20 or Cdh1 is likely to be bound and indeed the total binding interface spans more than just the the C-box motif. Therefore, relative to either Cdc20 or Cdh1 bound to Apc8, our C-box ligands offer a smaller binding interface that is unable to stabilise Apc8 to the same extent as either co-activating protein.

Table 6.7: Melting temperatures of Apc8 (1-591)-HiBiT in transfected HEK 293T lysates, with C-box ligands.

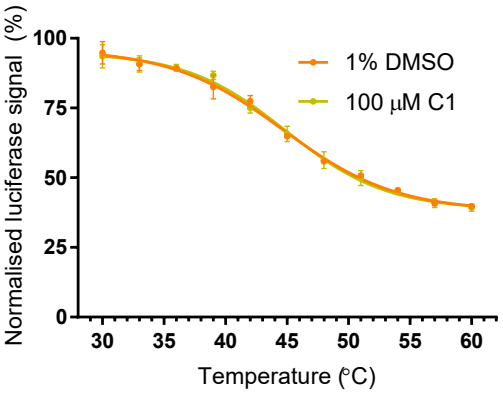
Ligand	Melting temperature (°C)
1% DMSO	44.3 ± 0.3
100 μM TAME	44.1 ± 0.4
100 μM C1	44.4 ± 0.6
100 μM C5	44.6 ± 0.2
100 μM C6e	42.0 ± 0.5
100 μM C6l	41.9 ± 1.0



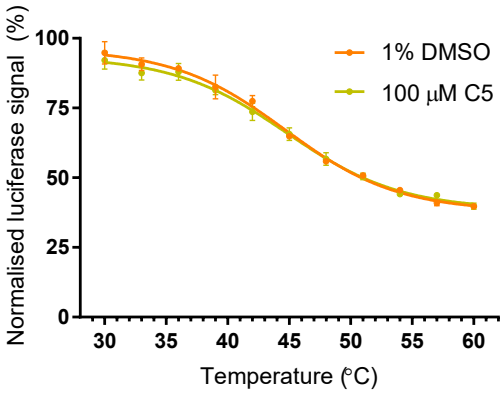
(a) Vehicle control



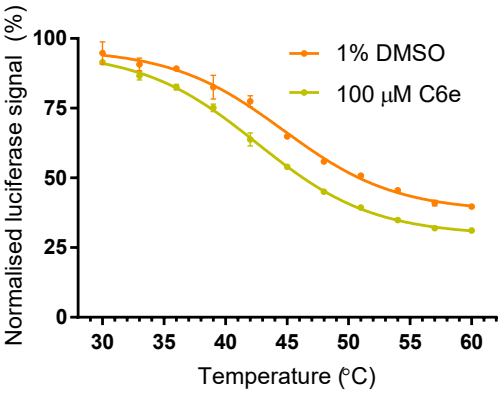
(b) TAME



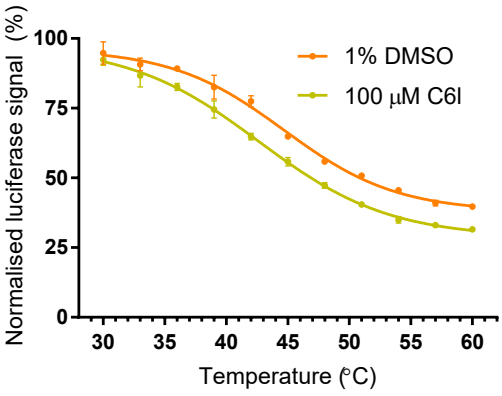
(c) C1: Ac-GDRYIPSR-NH_2



(d) C5: Ac-GDAAIPSA-NH_2



(e) C6e: Ac-S_5DRYIPS_5R-NH_2



(f) C6l: Ac-S_5DRYIPS_5R-NH_2

Figure 6.14: Thermal melt curves of exogenous Apc8 (1-591)-HiBiT in the presence of vehicle control (1% DMSO), TAME, C1, C5, C6e and C6l. (a) Impact of centrifugation as determined by comparing centrifuged (lime-green) and non-centrifuged (orange) samples of 1% DMSO control. Ligand-induced destabilisation of Apc8 using (b) 100 µM TAME, (c) 100 µM C1, (d) 100 µM C5, (e) 100 µM C6e or (f) 100 µM C6l. All samples are overlaid with the 1% DMSO control. Data represent the average of triplicate experiments.

6.3.7 *In vitro* ubiquitination assays

Binding of a ligand to its target receptor is merely a starting point to generating an inhibitory molecule. The ligand of interest must also be able to produce a functional effect as a result of receptor binding. Until this point, characterisation of novel D-box and C-box ligands has revolved around investigating their binding affinities to the target protein and their abilities to engage their target in a cellular context using CETSA assays. However, determining whether these ligands can behave as functional inhibitors of the APC/C is far more complicated as a matter, owing to the unique and complex biology that the APC/C performs and is indeed regulated by. To make matters more challenging in this regard, functional output of cell cycle control protein would ideally be investigated in living cells. However, peptide ligands on their own are typically impermeable to the cell membrane due to their hydrophilic nature, which was discussed previously in Section 1.2.1. To overcome this and also vastly simplify the complex nature of the APC/C function that occurs in a cellular context, *in vitro* ubiquitination assays that had previously been established by the Barford laboratory were employed to investigate the inhibitory action of the lead ligands (Figure 6.15).

The first iteration of this experiment utilised very high concentrations in the millimolar range since previous characterisation of TAME could only achieve inhibitory effects at concentrations in this range (Figure 6.15a) [28]. Indeed, poor inhibitory activity of TAME was observed in this experiment, whereby only marginal effect is observed at 3 mM concentrations. In stark comparison, both D-box peptides are significantly more potent at inhibiting Cyclin B1 poly-ubiquitination at the same concentrations. At 300 μM concentrations of D20 and D21, band intensity of each band in the poly-ubiquitinated Cyclin B1 ladder is less intense compared to the vehicle control. At the highest concentrations of 3 mM, the ladder intensity is further reduced, in addition to the extent of poly-ubiquitin laddering where only 2-3 bands in the ladder can be observed. However, it should be noted that TAME does not serve as the correct positive control molecule in this experiment due to its mechanism to deplete Cdc20 from the APC/C, whereas D-box peptides directly compete for the D-box degron binding site of Cdc20. As such, it was imperative to repeat these experiments with the appropriate controls. In addition, such high concentrations of ligands are physiologically irrelevant and as such, lower concentrations were used in further experiments.

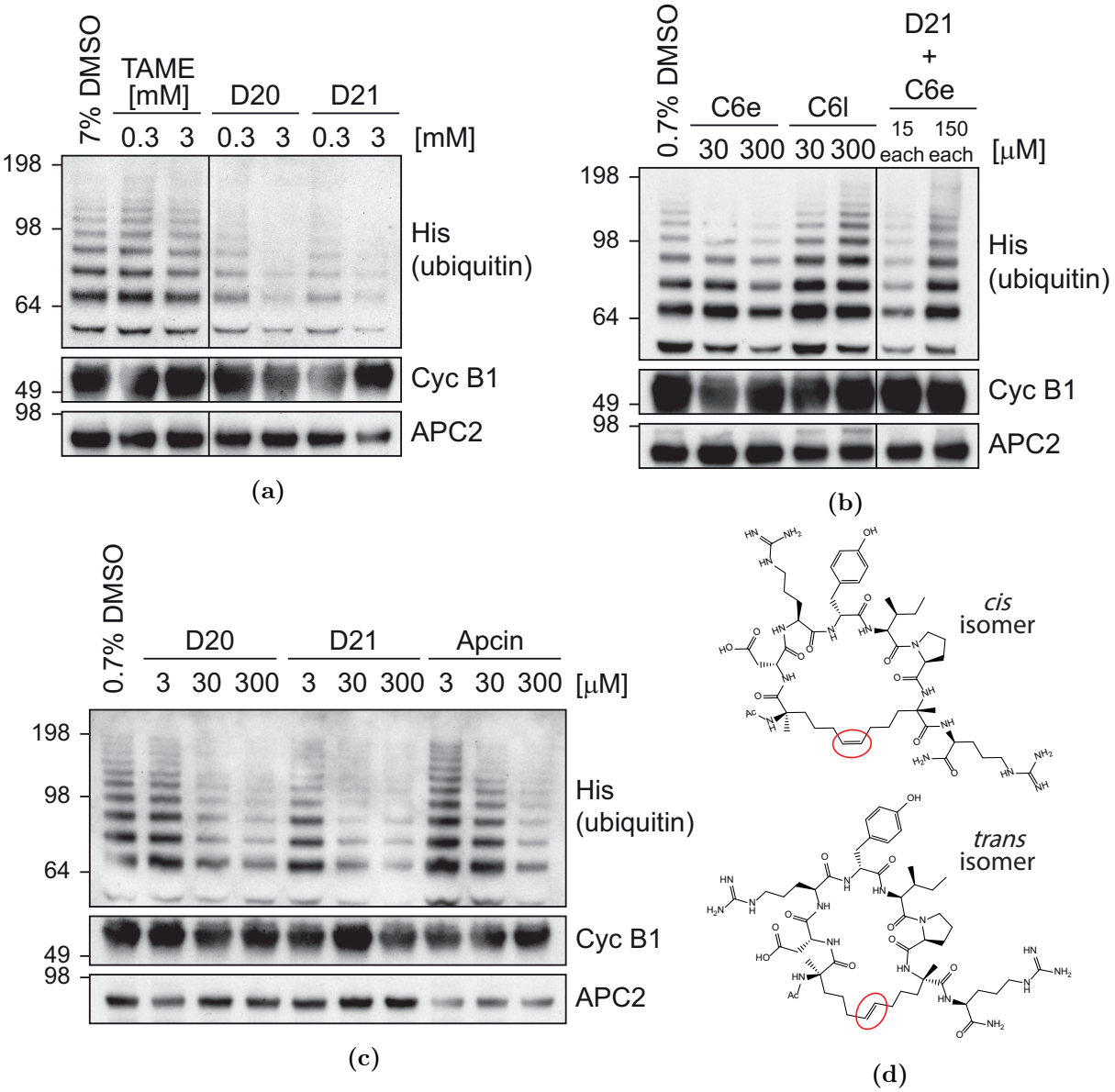


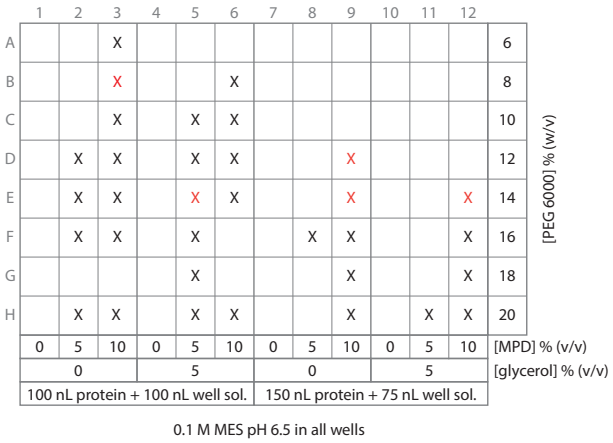
Figure 6.15: Ubiquitination assays probing Cyclin B1 poly-ubiquitination in the presence or absence of TAME, Apcin, D-box peptides and C-box hydrocarbon stapled peptides at varying concentrations or vehicle control. (a) Preliminary experiment probing TAME, D20 and D21. (b) Experiments to probe the C-box stapled peptides and a combination dosage with C6e and D21. (c) Experiments to comparing D20 and D21 against Apcin. Cyclin B1 and Apc2 serve as loading controls for each experiment. (d) Structures for C6e and C6l, which vary by *cis* and *trans* isomerisation at the carbon-carbon double bond of the hydrocarbon staple (circled in red) Further information of each peptide is described in Appendix B.

This pilot experiment was followed up with two experiments. Since the C-box stapled peptides produced a destabilising effect in the CETSA experiments (Figures 6.14e and 6.14f), it was believed that this may translate into a functional change of APC/C activity. The two ligands were tested at two concentrations of 30 μM and 300 μM, where both ligands produced surprisingly opposing functional effects. The stapled peptide, C6e, produces an apparent inhibitory effect on Cyclin B1 ubiquitination in comparison to the vehicle control lane (0.7% DMSO), which supports the hypothesis that Cdc20 could be displaced by the ligand as initially designed. In contrast, C6l produces the opposing effect, where increasing concentrations of ligand appear to further activate the APC/C and induce greater amounts of Cyclin B1 poly-ubiquitination compared to the vehicle control (Figure 6.15b). Given the high structural similarity of the two stapled peptides, which only differ by the *cis* and *trans* isomers at the carbon-carbon double bond of the hydrocarbon staple linker (Figure 6.15d), this result was highly surprising. In addition to single ligand dosing, as was done for all ligands, it was hypothesised that dual inhibitor action may result in a synergistic effect, as was observed by Sackton et al. [195]. When co-drugging with C6e and D21, which both produce inhibitory function in these assays, we observed quite the opposite when drugged in conjunction. Although at 15 μM of both ligands, inhibitory function is observed compared to the vehicle control, increasing of both doses 10-fold to 150 μM results in apparent lack and perhaps even further activation of APC/C ubiquitination.

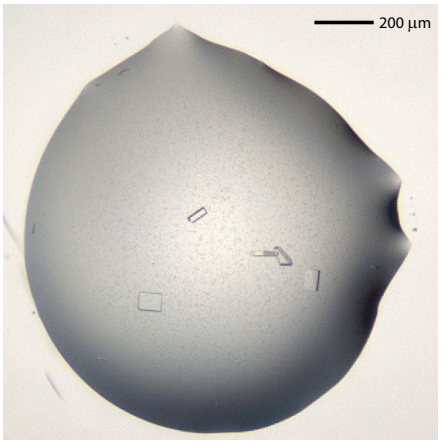
Lastly, we conducted experiments to directly compare the lead D-box peptides, D20 and D21 against the small molecule, Apcin. Although Apcin had not been previously assayed by the Barford laboratory, a dose-dependent inhibitory profile for the ligand was expected as seen in Figure 6.15c. Both D20 and D21 also produce similar inhibitory profiles for the 300 μM concentrations between Figures 6.15a and 6.15c, which provides confidence in the results observed. Curiously, despite the measured affinities of each ligand that were determined against Cdc20 in the SPR assays (Tables 6.2 and 6.4), which indicate that D21 and Apcin are very similar in affinity, and D20 is approximately 2-fold weaker than Apcin and D21, both peptide ligands appear to produce a greater inhibitory response, which can be most evidenced at the 30 μM concentration of each ligand. This result was highly intriguing and provides evidence that ligand affinity does not necessarily equate to a good inhibitor.

6.3.8 X-ray crystal structures of D-box peptides bound to Cdc20

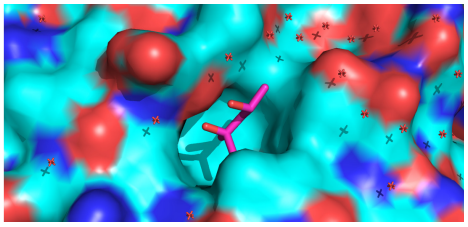
To obtain a more detailed understanding of the interaction between D-box peptides and Cdc20, structural information was thought to be necessary. However, due to the production of recombinant Cdc20 being a limiting factor, only one crystallisation grid screen was prepared and the Cdc20 concentration of 1.9 mg/mL was significantly lower than previously reported (7-8 mg/mL [193] and 4-5 mg/mL [195]). Despite this apparent disadvantage, cuboidal crystals were observed in multiple wells that grew to maximum size within 2-3 days of incubation at 20°C (Figures 6.16a and 6.16b). This surprising success elicited a serious emotional perturbation, the symptoms of which were excessive dancing and singing (of the human, not Cdc20). The MPD molecule contained in the mother liquor binds in the leucine-pocket of the D-box-binding site (Figure 6.16c) [193]. Therefore, it was thought unlikely that crystals would contain any D21 peptide even though it was used for concentrating the protein stock originally. Indeed, crystals that were not soaked in peptide did not have D21 bound.



(a)



(b)



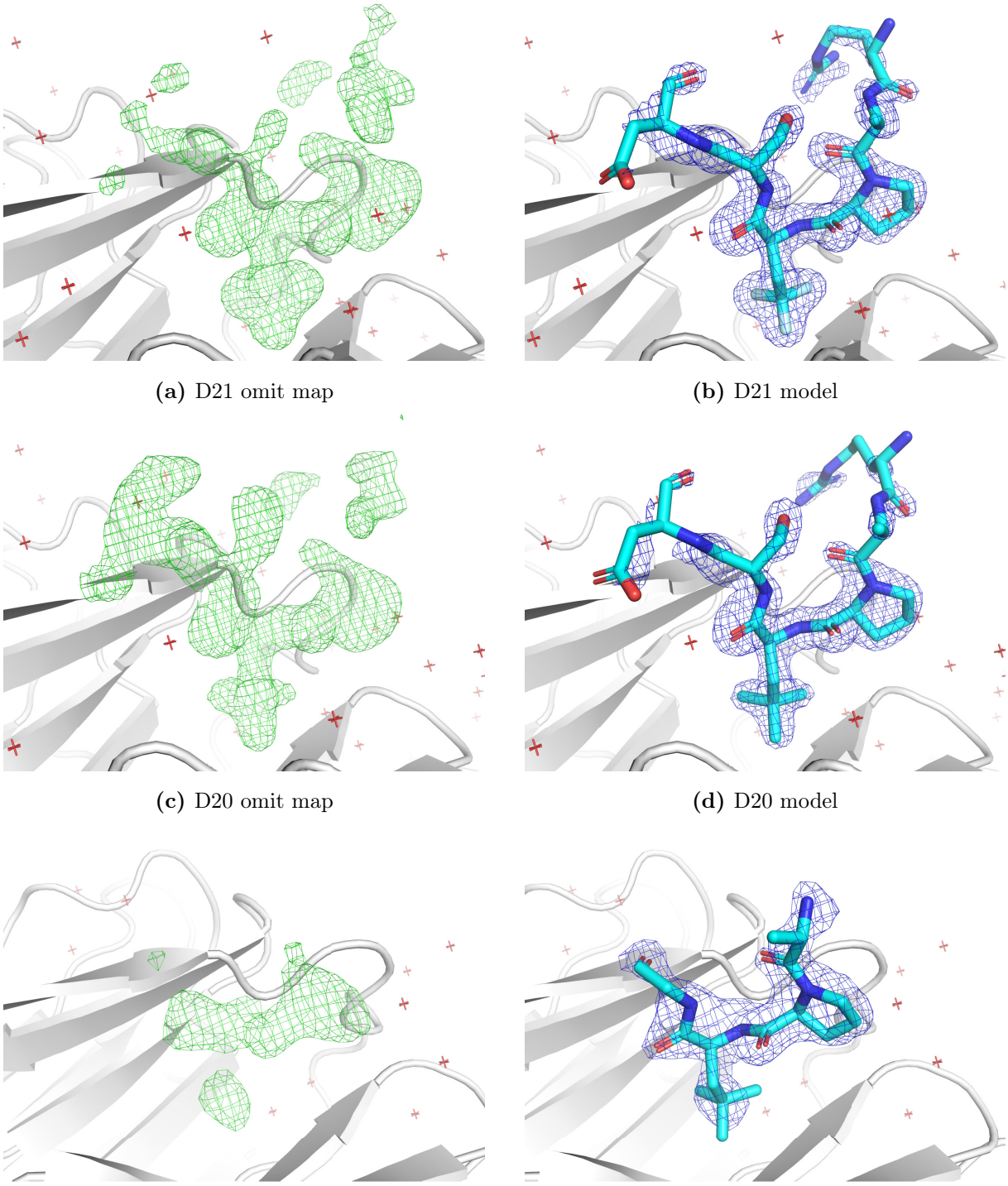
(c)

Figure 6.16: (a) Crystallisation hits are marked by a cross (X) in their corresponding wells. Red crosses (X) indicate wells containing crystals that were sent for data collection. (b) A representative image of cuboidal Cdc20 (161-477) crystals in drop E9, grown in 0.1 M MES pH 6.5, 14% (w/v) PEG 6000, 10% (v/v) MPD, 2:1 protein to well solution. A scale bar of 200 µm is shown as a reference for crystal size. (c) MPD binds the D-box binding site (PDB:4GGC) [193].

Crystals were soaked in solutions containing four D-box peptides: D21, D20, D7 and D19. Since, D7 and D19 are relatively weak ligands in the field of drug discovery, with dissociation constants in the low micromolar range, soaking was performed using these peptides at higher concentrations. The crystals did not appear to dissolve in the soaking conditions over the 4-hour incubation period. Structures containing D-box peptides were solved for the three highest-affinity D-box peptides (D21, D20 and D7) and the corresponding structure refinement statistics are summarised in Table 6.8. Densities for the D19 peptide were very poor, and it was not possible to build meaningful models. All structures solved, were found in the space group $P2_1$, as was observed previously [193]. Interestingly, the observed density in the omit maps (Figures 6.17a, 6.17c and 6.17e) correlates with ligand affinity. Due to the packing geometry, in which the next asymmetric unit is observed in very close proximity to the peptide binding site, densities for the C-terminal three residues (I[T/S]N-NH_2) of each peptide ligand could not be observed.

Table 6.8: Data collection, phasing and BUSTER refinement statistics for Cdc20-D-box structures.

Parameters and statistics	Crystal		
	Cdc20-D21	Cdc20-D20	Cdc20-D7
Data collection			
Space group	$P2_1$	$P2_1$	$P2_1$
Unit cell, a, b, c (Å),	35.49, 87.55, 48.57	35.34 87.34 48.51	35.00 86.87 48.03
α, β, γ (°)	90.00, 109.71, 90.00	90.00 110.14 90.00	90.00 109.60 90.00
Resolution range, Å	45.73 - 1.51 (1.66 - 1.51)	45.55 - 1.46 (1.60 - 1.46)	45.24 - 1.92 (2.09 - 1.92)
Total reflections	162715 (6076)	184416 (10465)	76640 (3657)
Unique reflections	32820 (1641)	35450 (1772)	14979 (750)
Multiplicity	5.0 (3.7)	5.2 (5.9)	5.1 (4.9)
Completeness (spherical), %	74.7 (14.5)	74.5 (16.0)	71.8 (15.4)
Completeness (ellipsoidal), %	90.4 (41.9)	92.9 (60.1)	90.5 (57.6)
I/σI	13.1 (1.6)	13.6 (1.4)	7.6 (1.5)
R_{merge}	0.050 (0.598)	0.048 (1.000)	0.132 (1.161)
$CC_{1/2}$	0.999 (0.714)	0.999 (0.635)	0.996 (0.536)
Refinement			
R_{work}/R_{free}, %	0.175/0.194	0.165/0.184	0.207/0.231
Unique reflections used	32820	35450	14966
R.m.s deviations:			
bond lengths, Å	0.008	0.008	0.008
bond angles, °	1.01	1.02	1.00
Ramachandran analysis:			
Favoured, %	97.10	98.06	97.33
Allowed, %	2.58	1.61	2.33
Outliers, %	0.32	0.32	0.33
Number of atoms			
(average B-factor, $Å^2$):			
Protein	2452 (24.31)	2434 (25.80)	2340 (29.68)
Solvent	236 (39.17)	240 (44.08)	72 (32.96)
Peptide	46 (37.72)	46 (46.88)	25 (48.49)
Mean/Wilson B-factor, $Å^2$	25.8/22.9	27.8/23.9	30.0/27.8



(a) D21 omit map

(b) D21 model

(c) D20 omit map

(d) D20 model

(e) D7 omit map

(f) D7 model

Figure 6.17: X-ray crystal structure data of D-box peptides in complex with Cdc20. (Left panels) Omit maps of refined structures showing Fo-Fc map (green mesh) contoured at 2.5σ for peptides (a) D21, (c) D20 and (e) D7. (Right panels) Models of peptides built into the electron density map showing 2Fo-Fc map (blue mesh) contoured at 1.0σ for peptides (b) D21, (d) D20 and (f) D7. Peptides in stick representation are coloured in cyan. Cdc20 represented in cartoon is shown in white.

Next, the structures were compared to previously published Cdc20 structures (Figure 6.18) [193, 195], resulting in an alignment with a negligible root mean squared deviation (rmsd) of 0.228 Å between all residues and atoms of the Cdc20 protein in PDB:4GGC and the Cdc20-D21 structure, proving that these Cdc20 structures are highly similar. Minor shifts were only observed around a loop comprising residues 340 - 344 and could be due to a the poor density and consequently modelling of these local residues. In the case of the Cdc20-D7 structure, it was not possible to build convincing residues for this region.

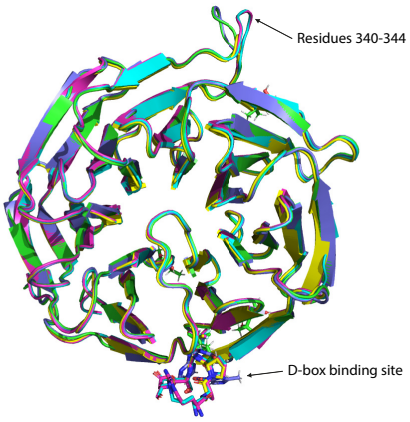
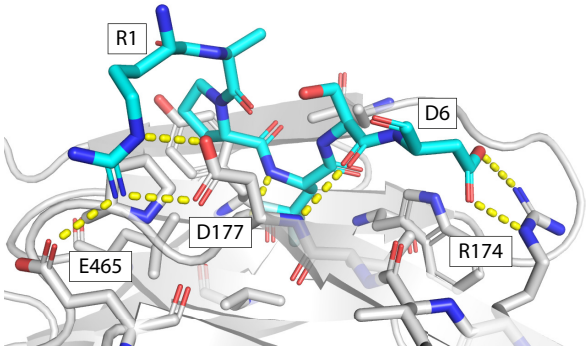


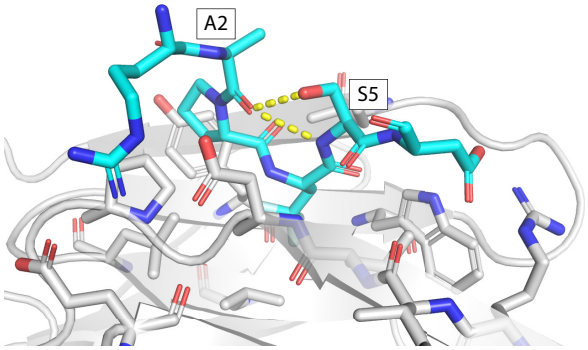
Figure 6.18: Alignment of published Cdc20 structures with the structures of Cdc20 bound to D-box peptides. The colouring is as follows: PDB:4GGC - green, PDB:4N14 - violet, Cdc20-D21 - cyan, Cdc20-D20 - magenta, and Cdc20-D7 - yellow. All water molecules were omitted for clarity.

In all three structures, peptides bound the canonical binding site and formed the same set of inter-molecular hydrogen bonds (H-bonds), as can be seen in Figures 6.19a, 6.19c and 6.19e. The guanidino group of R1 of peptides interacts through H-bonding with the carboxylic acid side chains of D177 and E465 of Cdc20. The nitrogen backbone atom of the F_3/C_3 non-natural amino acids also form a hydrogen bond with the carbonyl of D177. Additionally, the D177 nitrogen backbone atom forms an interaction with the carbonyl group of S5/G5 of the peptide. Lastly, D6 forms inter-molecular H-bonds with R174. This network of H-bonding appears to be identical to the H-bond network described previously for homologous residues in the Acm1-Cdh1 structure [165].

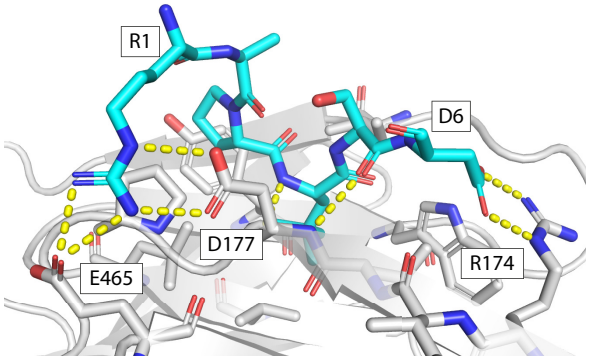
In addition to the inter-molecular H-bonding network, the incorporation of S5 in D21 and D20 instead of glycine results in an unique intra-molecular H-bond formed between the hydroxyl groups of the serine side chain with the carbonyl group of A2. A second intra molecular H-bond is formed between the nitrogen backbone atom of either residue at position 5 and the carbonyl of A2, although this is likely to be a weaker given its longer range (3-3.1 Å).



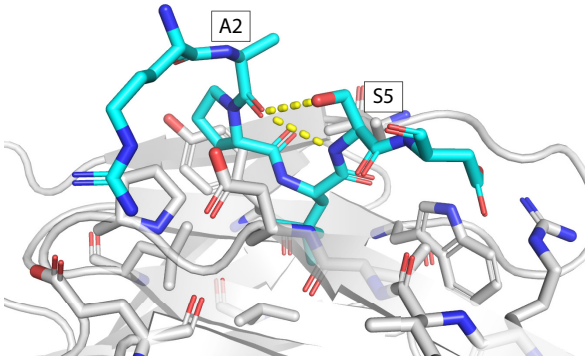
(a) D21 inter-H-bond



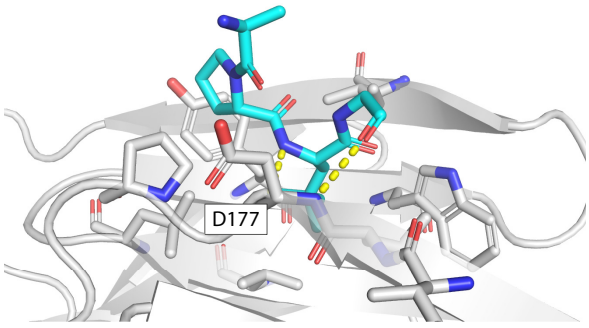
(b) D21 S5 intra-H-bond



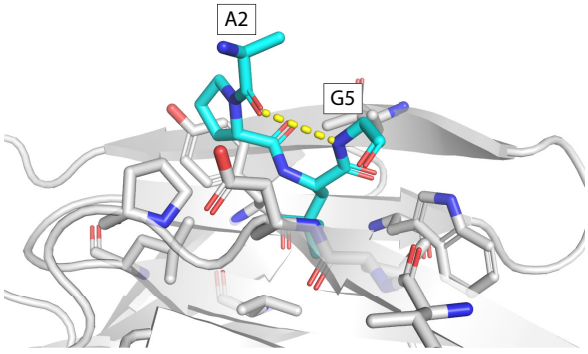
(c) D20 inter-H-bond



(d) D20 S5 intra-H-bond



(e) D7 inter-H-bond



(f) D7 intra H-bond

Figure 6.19: X-ray crystal structure data of D-box peptides in complex with Cdc20. Left panels: Inter-molecular hydrogen bonding (yellow-dashed lines) between Cdc20 (R174, D177, E465) and D-box peptides (a) D21, (c) D20 and (e) D7. Right panels: Intra-molecular hydrogen bonding (yellow-dashed lines) between position 5 amine group group and A2 carbonyl group for peptides (b) D21, (d) D20 and (f) D7. D21 and D20 contain serine at position 5, which contributes a second intramolecular H-bond between the serine hydroxyl group and A2 carbonyl group.

6.4 Discussion

Biophysical assay development

The development of assays to assess peptide binding were mindful of the limitations involved with producing purified Cdc20. We found DSF to be very efficient in terms of sample use, requiring a concentration of only 750 nM to produce sufficient signal from Sypro Orange binding upon the unfolding of the WD40 domain. Indeed, this often requires protein concentrations in the range of 5-20 μM to produce a sufficient signal-to-noise ratio. Although it is possible to derive affinity constants from DSF data, either by extrapolation of thermodynamic parameters [245] or by fitting isothermal data from a titration of ligand concentrations [246], these routes were not explored. Firstly, the method of extrapolating thermodynamic parameters requires defining the specific heat capacity of the protein, which can be experimentally determined using differential scanning calorimetry (DSC). Although, it is possible to estimate this value based on the molecular weight of the protein, this may result in a large and unknown error. Secondly, DSF was used only as a preliminary screening method, in which the titration was only performed over three ligand concentration points. This was sufficient to provide confidence that the change in melting temperature was due to the stabilisation of Cdc20 without using larger amounts of protein which would be required for a titration that can be used to determine an affinity.

The phenomenon that a given binding ligand can raise the T_m of a protein, even in vast excess of the apparent stoichiometry and exceeding the ligand's K_D to the target is one that is often observed [245, 247]. Although these authors have dedicated a significant amount of time in the development of mathematical models to describe differential scanning fluorimetry data, they do not answer why this phenomenon occurs.However, it is clear that this phenomenon is independent of the protein class, whether it is a reversibly folding protein or not and indeed what kind of ligand it may be. The only common characteristic is that the ligands are reversible binding partners to the protein target. Though detailed models and descriptions describing this could theoretically be pursued, it is certainly beyond the scope of this PhD. However, one way to briefly rationalise the phenomenon may be by thinking about the change in K_D of a ligand as a function of temperature. Indeed, with higher temperatures, 10-fold weaker binding affinities can be observed, between 6°C and 37°C [248]. Most physiological experiments will not exceed temperatures greater than 37°C, but here the goal is to fully induce thermal denaturation of the target protein. A key question in the context of thermal denaturation assays is how weak is the K_D at the T_m of the target protein. In addition to this, it may also be important to consider the kinetics of protein unfolding in relation to the kinetics of ligand association and dissociation and also as a function of temperature.

To further characterise each peptide, SPR experiments were performed. The Ni-NTA immobilisation strategy displayed poor activity of the immobilised Cdc20 protein. This may be due to the subsequent amine coupling required to ensure the RU_{ligand} remains sufficiently stable over the course of an experiment. It nevertheless proved worthwhile to develop a SAR for for peptides in the early stages of the project. Minimally biotinylated Cdc20 was significantly more active on the SA biosensor, observing ligand activities around 45% for D-box peptides. Although this activity is still far from 100%, it may be attributed to subjecting the Cdc20 protein to multiple freeze-thaw cycles between purification, biotinylation and SPR experiments. In addition, the data observed for Apcin, D21 and D20 would be sufficient for kinetic fit analysis, given less bulk RI shift upon valve switching at the beginning and end of the analyte association phases. However, this would likely require thorough cleaning of the Biacore instrument prior to experimentation and good sample preparation to avoid DMSO mismatching. Ideally, an orthogonal technique to measure binding affinities of peptides would have been developed, such as competition fluorescence polarisation assays. However, this technique uses high concentrations of protein dependent on the K_d of the fluorescent-labelled tracer peptide. Therefore, we were unable to develop a tracer peptide of sufficient affinity early in the project to aid our screening approach.

At present, no biophysical characterisation of Apcin binding to Cdc20 has been explored, except the CETSA and X-ray crystal structure datasets published by Sackton *et al.* [195]. However a review by Schapira *et al.* [249] claim Apcin to have a $K_D > 10$ μM, without providing further citation. This value is in fact an error that relates to the K_i of the ligand instead. The SPR data presented here show that the K_d is much closer to 500 nM, a 20-fold difference to the previously claimed value. Given this high affinity, Apcin was the ideal positive control molecule during the peptide screens in addition to providing an aim to design a peptide of at least as good an affinity if not better. Moreso, the development of robust DSF and SPR assays provide opportunities for further studies to develop inhibitors against all Cdc20 degron-binding sites, a key advancement for the scientific community.

Although assays to generate reliable data to build a SAR for D-box, KEN-box and ABBA motif peptides were developed, it was not as simple to develop a method for the assessment of C-box peptides. The target of the C-box peptide was primarily Apc8, with a potential to interact further with Apc3 due to strong homology of C-box and IR-tail binding pockets, respectively. Both proteins are essential components of the core APC/C structure [158] and therefore the proteins themselves (or truncated subdomains)

can not be purified without their interaction partners (communications with David Barford, unpublished). Therefore, CETSA assays were explored for Apc8, Apc3 and Cdc20, simultaneously, using the small molecules, Apcin and TAME as initial ligands in early experiments. The approach using quantitative western blotting was beneficial since the readout was from endogenous protein. However, the generation of data was cumbersome and prone to error at multiple stages. These errors could arise from unequal transfer efficiencies as well as the amount of endogenous Cdc20 not being sufficient to create a significant signal and resolution, as could be observed from the relatively small thermal shifts upon addition of the small molecules. The initial results also suggested that this method was unlikely to provide sufficient resolution to understand the SAR of D-box, KEN-box, ABBA motif or C-box peptides. However, a subset of the highest affinity D-box peptides were tested in a split-luciferase CETSA experiment, which was considerably higher throughput and generated data that were in support of the affinity trends observed by DSF and SPR. Yet, small differences in the melting temperatures of Cdc20 with 100 μM Apcin were detected between the split-luciferase and the western blot assays. This may be due to the significantly lower concentration of total protein and therefore lower concentrations of Cdc20 used in the split-luciferase CETSA (0.2 mg/mL total protein) compared to the Western blot (2.27 mg/mL total protein). Further experiments would be required to prove this hypothesis.

Similarly to split-luciferase CETSA assays for Cdc20-binding ligands, constructs were generated to assess ligand binding to Apc8 using the split-luciferase assay. It is believed this is the first semi-biophysical method to observe ligand binding to any core component of the APC/C. This provides a particular challenge since true physiological relevance requires Apc8 to be in complex with the APC/C. In this work, experiments to evidence this have not been explored due to a lack of time. Whilst it was possible to setup the split-luciferase CETSA assay for Apc8, attempts to do similar for the IR-tail-binding subunit, Apc3, were not successful. Preliminary experiments showed very low raw luminescence signal, only between 6-8-fold greater than the background signal. Comparatively, Apc8 and Cdc20 produce greater than 100-fold and up to 400-fold increased signal compared to the background, respectively. The reason that split-luciferase based Apc3 CETSA assay development was unsuccessful in our hands could be two-fold. Firstly, the HiBiT-tag, irrespective of which terminus it is located, could be buried within the native protein and therefore unable to complement the LgBiT protein to reconstitute the NanoLuc. Secondly, and perhaps the more likely case is that incorporation of the HiBiT-tag on either end hinders the incorporation of the Apc3 subunit into the larger APC/C complex. Based on our previous communications with Prof. David Barford, this cellular protein is likely to be highly unstable and could therefore be rapidly degraded by the cell as it is expressed.

The CETSA data suggests that peptides can bind full-length Cdc20 and Apc8 within a cellular context. However, CETSA experiments can be performed in live cells and whole tissues, but since it was unlikely that the peptides would traverse the cell membrane and subsequently reach their intracellular targets, experiments were limited to cell lysates and therefore are not a true representation of a living cell. Furthermore, the peptides are likely to be cleaved by broad specificity cellular proteases, so the inclusion of protease inhibitors were key for retaining intact ligand for the course of the experiment.

Structure-activity-relationships of the D-box peptide

Through iterative design, synthesis and characterisation of D-box peptides it was possible to deconvolve the structural determinants corresponding to affinity. Early in the development, isoleucine at position seven was identified as the optimal residue of three aliphatic side chain amino acids (valine, isoleucine and leucine). Following this, proline at position 3 was believed to provide optimal peptide-backbone geometry for the turn between residues 2-5. Subsequent incorporation of S5 and T8 from D1 into the D4 backbone generated the lead natural amino acid-containing D-box peptide (D19, Ac-RAPLSDITN-NH_2) with an affinity of 5.87 ± 0.11 μM. Although the D19 peptide did not exhibit sufficient density *in crystallo*, the X-ray datasets of D20- and D7-bound Cdc20 provide some insight into why S5 is preferred over G5. The formation of a hydrogen bond between the hydroxyl side chain of serine and carbonyl group of alanine provides an intra-molecular stabilisation of the bound peptide. One would therefore expect that this is only important in the bound state of the ligand and so may have an affect on the k_{off} rate, rather than the k_{on} rate. Although this could be explored by SPR, the dataset collected was not of sufficient quality for kinetic fitting.

The addition of C_3 or F_3 non-natural amino acids at position 4 improved the peptide affinity 6- or 11-fold, respectively, resulting in affinity of peptide D21 (Ac-RAPF_3SDITN-NH_2, 524 ± 14 nM) that is comparable to Apcin (418 ± 53 nM). It was highly unfortunate that the Fmoc-(*S*)-2-amino-4,4,4- trichlorobutanoic acid could not be synthesised. Due to the evidence from the binding of Apcin, which contains a CCl_3 group, we expect the affinity of a peptide with the sequence (Ac-RAPCl_3SDITN-NH_2) to further improve and perhaps reach the double-digit nM range. Indeed, molecules with binding affinities in this range become successful drug candidates.

In vitro ubiquitination assays

The use of functional assays to determine whether both D-box and C-box ligands could inhibit the E3 ubiquitin ligase function of the APC/C was invaluable to this project.

With these few experiments, it was demonstrated that the extent of poly-ubiquitination of a bait substrate, Cyclin B1, could be greatly reduced in the presence of D-box peptides D20 and D21 at concentrations ranging from 30 μM up to 3 mM. Most notably, both peptides appear to provide greater reduction of Cyclin B1 ubiquitination by the APC/C compared to Apcin at equivalent concentrations. Since both D20 and D21 are less potent binders than Apcin, this suggests a role beyond the standard affinity dogma of inhibition. Indeed, D-box degron recognition by Cdc20 and Cdh1 is deemed to require an Arginine at position 1 of the degron sequence (Figure 5.4, [165]) for effective recognition. However, Apcin as a ligand primarily competes for the Leucine binding pocket at position 4 of the D-box degron (Figure 5.7, [195]). This begs the question whether D-box degrons contained within Cyclin B1, in our assays can still interact and be poly-ubiquitinated by the APC/C despite Apcin being bound to Cdc20. In contrast, since our D-box peptides contain the Arginine residue at position 1, and indeed even more of the designated D-box degron towards the C-terminus, it is believed that any transient and/ or weak interactions of Cyclin B1 with Cdc20 in the presence of Apcin are not feasible with D20 or D21 given their larger extended binding interfaces. This is also particularly interesting since more recently following the submission of this thesis, a further publication relating to Apcin's mechanism of action were published by Randall King's laboratory [199]. Briefly, given the right cellular conditions, where the spindle assembly checkpoint activity is high, Apcin can increase the rate of mitotic exit thus producing the opposite desired phenotypic effect. Whether the novel D-box peptides would act similarly is a question that would still need to be answered.

In addition to the D-box ligands, stapled C-box peptides were also tested in these assays. These two ligands, C6e and C6l, corresponding to the early and later isomers (determined by analytical HPLC) of the hydrocarbon staple are particularly interesting due to their apparent opposing functions in the *in vitro* ubiquitination assay. C6e's dose-dependent inhibition on Cyclin B1 ubiquitination fits the model of an inhibitory peptide that acts to deplete Cdc20 from the APC/C, by competing for the C-box binding site of Apc8. On the other hand, C6l, produces the opposing effect, whereby Cyclin B1 ubiquitination is vastly increased in a dose-dependent manner. Why exactly this occurs is currently unknown and would require extensive characterisation using multiple techniques to determine its mechanism of action. When considering that these two peptides only differ by the orientation of the carbon-carbon double bond in the *cis* or *trans* isomer, these results are even more surprising and fascinating.

Alongside looking at C6e and C6l as inhibitors in their own right, it was hypothesised that co-treatment of our lead D-box peptide, D21, and the stapled peptide C6e, that produces an inhibitory effect as a singular treatment could produce a synergistic effect on

Cyclin B1 ubiquitination. This principle had been established with the TAME and it's prodrug analogue proTAME with Apcin by Sackton et al. [195]. Initially, it was envisaged that since both D21 and C6e outperform their small molecule counterparts, an even greater reduction of Cyclin B1 ubiquitination would be observed. However, as observed in Figure 6.15b, the opposite was evident. Although at the lower concentrations tested, the overall ubiquitination is less than the vehicle control, increased drug concentrations result in more ubiquitination, akin to the observations with C6l. Why this is occurs, is still unknown and is highly puzzling. However, it very succinctly demonstrates the difficulty and complexity of effectively inhibiting large complexes, such as the APC/C.

X-ray crystal structures of D-box peptides bound to Cdc20

These datasets presented in this chapter are the first known structures of designed D-box peptide inhibitors bound to Cdc20. Although the overall molecular architecture of binding is highly similar to the D-box peptide in the Acm1-Cdh1 structure [165], subtle differences can be detected. We observed very strong evidence for the binding of the F_3 and C_3 amino acids in the leucine-binding pocket. In addition, the data suggest why serine at position 5 may improve the affinity of the peptide due to geometry of hydrogen bonding between the hydroxyl group of serine to the carbonyl group of alanine at position 2. Unfortunately, residues 7-9 remain unresolved due to steric clashing between the peptide C-terminus and the protein in the neighbouring asymmetric unit. Since the peptide can only bind through its six N-terminal residues, the binding affinities of the whole peptide in the context of the crystal lattice are likely to be impaired and this may be the reason why it was not possible to build convincing models for D19. Currently, all Cdc20 structures (without additional proteins) are solved in the space group $P2_1$ or $P2_12_12_1$ [193]. However, even in the alternative space group, Cdc20 from the neighbouring asymmetric unit partially occludes the D-box binding area. Unless the protein is crystallised in a different space groups, it is unlikely that the entire peptide will ever be resolved.

In addition to the asymmetric unit clashing with the C-terminus of the peptides, there is also another possible explanation to the lack of density for these residues. It has been proposed that one of the core subunits of the APC/C, Apc10, acts as a co-receptor for D-box degrons contained within substrate proteins [250–252]. Da Fonseca et al. show an overall shift of Cdh1 towards the Apc10 subunit in the presence of substrates containing a D-box degron. More recent higher resolution cryo-EM structures from the Barford laboratory indicate that the C-terminal residues comprising the D-box degron are directly involved in forming the Apc10 co-receptor, as shown in Figure 6.20. This raises the question whether the final two residues, i.e. TN-NH_2 could be observed without hindrance from the asymmetric unit.

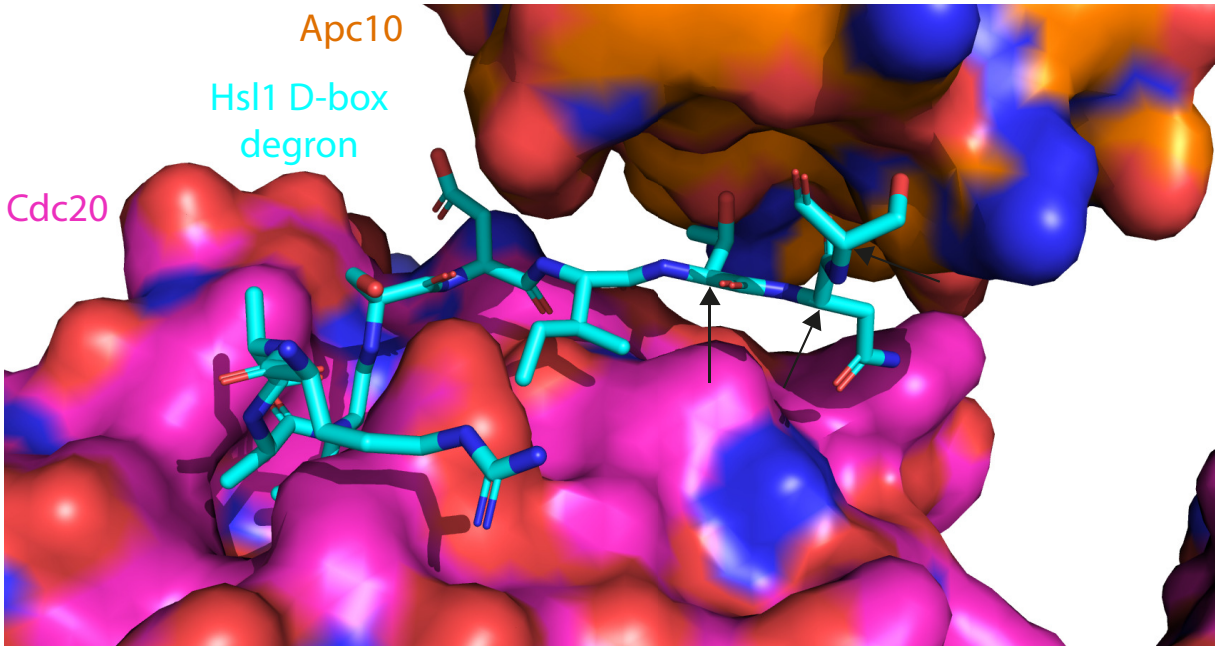


Figure 6.20: Cryo-EM model of Hsl1 D-box peptide (cyan, stick representation) in complex with Cdc20 (magenta) and Apc10 (orange). Black arrows refer to the amino acids that interact with Apc10 at the C-terminus of the D-box degron sequence. PDB ID: 5G04, [28].

Following the defence of this thesis, and in preparation of the data for publication, it was realised that an undesirably large R_{work}/R_{free} gap had emerged in the Cdc20-D7 structure. In addition, the electron density maps for this structure were of less quality compared to the other two, taking into account the difference in resolution. Following investigation, it was found that this gap would occur immediately after the first round of refinement in Refmac 5.0, following the molecular replacement solution by Phaser. Although it was never understood why this occurred, it was decided to redo the refinement using BUSTER (Global Phasing), as suggested by Dr. Marianne Schimpl at AstraZeneca. Refinement in BUSTER resulted in better map quality, containing less large areas of negative density. In addition, the R_{work}/R_{free} gap closed up towards acceptable values for publication and altogether there was better confidence in the modelled peptide residues. For consistency between the three structures, it was decided that all datasets would be refined using BUSTER instead. In all cases, the R_{work}/R_{free} gap was smaller, the root mean square bond lengths and bond angles were improved and the overall quality of maps were better as colleagues at AstraZeneca had often observed with BUSTER refinements.

Previous attempts by Tian et al. to co-crystallise a D-box peptide derived from Securin

(CDAPPALPKAT**RKALGTVN**RATEKS) were noted as being unsuccessful despite obtaining high resolution data (1.35 Å). Note that the amino acids pertinent to the D-box degron are highlighted in bold. With the blend of knowledge gained from various aspects of this project, it is possible to understand why this might have occurred in their hands. Firstly, to dissect the structural determinants of D-box degron binding, the presence of Alanine at position 3 results in an approximately 3-fold weaker ligand compared to Proline at the equivalent position. It is also noted that Glycine at position 5 was weaker due to the loss of an intramolecular hydrogen-bond compared to its serine-containing counterpart. Position 6 also has Threonine in place of Aspartic acid in our synthesised ligands, which produced additional salt bridges between R174 of Cdc20. While threonine could potentially form hydrogen bonds via the hydroxyl group, the overall length of the side chain is also different and it therefore may not be within close proximity to form this bond. Lastly, position 7 contains a Valine residue, which was discovered early in our biophysical characterisation to be the worst of the three aliphatic residues for binding to Cdc20, approximately 2-fold weaker than an isoleucine-containing peptide. In addition, soaking experiments of the D21, D20, D7 and D19 ligands provided evidence of decreasing ligand density, which clearly correlated to the overall affinity of the ligand, as measured by SPR. The density around the expected binding site of the D19 peptide was so poor that it was not possible to build even a single residue of the peptide. The gained understanding of an approximate binary affinity for the securin D-box peptide used by Tian et al. is almost certainly in the range of 10-100 μM, if not even weaker. Given that D19 binds with 5.87 ± 0.11 μM, and no credible ligand density was observed for this peptide, it is not surprising that ligand density was absent using peptides that are undoubtedly weaker affinity in their attempts.

Altogether, these datasets provide evidence for the molecular architecture of binding between a designed D-box peptide and Cdc20. Originally, functional experiments were out of the scope of the initial project due to time constraints, lack of cell permeability/delivery methods and limited in-house expertise in this area of cell biology. However, the data presented here can now form a basis for further D-box peptide development and their characterisation in functional assays, such as the inhibition of the E3 ubiquitin ligase APC/C and consequently mitotic exit.

Chapter 7

Final conclusions and future work

The work described in this PhD thesis focuses on the development of peptide inhibitors against two tandem-repeat proteins whose over-expression is associated with cancer. For both targets, small molecule therapeutics are of limited efficacy, or have not yet been identified, and therefore peptide-based alternatives were explored using a rational structure-based design approach.

7.0.1 Development of Tankyrase inhibitors

The sub-project conducted in the first year of the PhD focused on the improvement of peptide inhibitors of the substrate-binding ARC domains of Tankyrase that were previously developed in the Itzhaki lab [102]. The original peptides had sub-micromolar affinities and, after conjugation with cell-penetrating peptides, they were shown to be able to inhibit tankyrase, however their IC_{50} values were not sufficiently efficacious for further drug development. Therefore, our primary goal was to further increase their binding affinities and thereby improve their cellular efficacy utilising the same delivery approach with CPPs. First an expanded suite of biophysical assays was developed, including FP, DSF, SPR and ITC. However, we found that SPR did not provide suitable data quality due to unknown reasons. Multiple approaches were explored to improve peptide affinity to the TNKS ARC domains. First, point mutations at position 5 of the peptide sequence were introduced to exploit an adjacent pocket of the surface that was observed in ARC2 and ARC5 domains compared to ARC1 and ARC4 domains. Although this approach did not yield improved affinities, this highlighted a degree of previously unknown specificity between the different ARC domains, where preferential, but weak binding was observed to TNKS1 ARC2-3 compared to TNKS2 ARC4. In addition, we explored a novel macrocycle design that yielded peptide T1-1, which bound with 3.41 $\pm$ 0.08 μM affinity to TNKS2 ARC4. Unfortunately, attempts to co-crystallise the TNKS2 ARC4-T1-1 complex were unsuccessful. Therefore, without an understanding of the macrocycle pose and why it re-

duced the affinity compared to the consensus sequence, no further structure-based design was possible.

Although peptides with higher affinities were not obtained as a result of these efforts, there is considerable scope to expand the project to include alternative methods. Although a consensus binding motif was described by Guettler *et al.* [131], it may not necessarily be the highest affinity ARC-binding peptide. First, library techniques such as phage, mRNA or ribosome display may produce a higher affinity peptide than the current consensus sequence, which was developed through matrix screen approach [131]. Second, different macrocyclisation chemistries could be explored, which are likely to sample alternative chemical space resulting in a more potent ligand. Third, the ability of Tankyrases to bind their substrates through four independent ARC domains allows the development of multi-valent inhibitors as a strategy increase cellular efficacy (ongoing work in the Itzhaki lab and publications to be submitted). Fourth, peptides could be designed that go beyond traditional competitive inhibition. For example, covalent peptide inhibitors can exploit the reactivity of lysine or cysteine side-chains surrounding the substrate-binding site. Lastly, peptide inhibitors of Tankyrase can potentially be used in combination with degradation-targeting moieties to harness the ubiquitin-proteasome system (UPS) and drive the degradation of Tankyrase to reduce its oncogenic activity of the various pathways that it can act upon such as the β-catenin/WNT signalling pathway [112]. Such work is currently underway in the Itzhaki lab.

7.0.2 Development of APC/C^{Cdc20} inhibitors

The greater focus of this thesis has revolved around the development of peptide inhibitors to target the oncogenic activity of the Anaphase Promoting Complex/Cyclosome with the Cdc20 co-activating protein. Cdc20 in stark contrast to it's homolog protein Cdh1, is known to be an oncoprotein. Unfortunately, the number of drug discovery campaigns has been limited to a high-throughput screen of over 100,000 small molecule compounds by Verma et al. [197]. In this screen, a number of compounds were found to stabilise the Cyclin B1-luciferase reporter used, but only three of these hits were found to be specific competitive inhibitors against the APC/C^{Cdc20} [194, 195, 198]. This equates to approximately 0.003% hit success rate. As such, we focused our studies on the development of peptide-based inhibitors to target the APC/C^{Cdc20}'s oncogenic activity. Initially, four peptide series were designed, consisting of peptides based on the three degrons (D-box, KEN-box and ABBA motif) with the fourth being based on the C-box motif required for Cdc20 to bind the Apc8 subunit of the APC/C.

We began our studies by attempting to setup a robust system in which to quantify the binding affinity of designed and synthesised ligands to Cdc20. For this, recombinant Cdc20 comprising the WD40 domain had to be expressed in baculovirus expressions systems and purified using a three-step chromatography protocol based on procedures described by [193, 195]. Obtaining pure Cdc20 presented a significant challenge to the project as a result of problems at multiple stages including expression, solubility and purification. Therefore, biophysical techniques were carefully selected that would be sufficiently sensitive with small sample requirements. DSF relies on a signal change when a fluorescent reporter molecule binds to hydrophobic residues that are exposed as the protein unfolds. The WD40 domain of Cdc20 contains high percentage of hydrophobic amino acids, which resulted in a large signal change at lower protein concentrations than are commonly used in this type of experiment. DSF proved to be an ideal method to screen a large number of peptide ligands relatively rapidly. This allowed us to initially screen our synthesised ligands to be ranked in order of affinity, of which the best were selected for further characterisation. The binding affinities were then quantified by SPR using biotinylated Cdc20 protein, and the results were in agreement with the rankings obtained by DSF. While this approach worked with probing binding to Cdc20, developing a method to explore the binding of the C-box peptides to Apc3 and Apc8 was not straightforward. We first utilised CETSA with endogenous protein to probe binding of the published small molecules Apcin and TAME to Cdc20 and Apc3/Apc8, respectively. However, using quantitative western blotting to detect thermal stabilisation of endogenous proteins was low-throughput and was not sufficiently sensitive. We therefore invested into

developing a more high-throughput method in 384-well plate format for CETSA assays using split-luciferase detection rather than western blotting. This allowed us to develop the first high-throughput experiments for screening Cdc20 and Apc8-targeting ligands, which can be utilised downstream by the drug discovery community.

Initial sets of the four peptide series, including a subset of D-box, KEN-box and C-box hydrocarbon-stapled derivatives, were synthesised in Singapore with the help of Dr. Fernando Ferrer (A*STAR). The designs were supported with modelling and MD simulations conducted by Dr. Yaw Sing Tan (A*STAR). Early in the biophysical analysis, the KEN-box and ABBA motif peptides were not observed to bind to Cdc20, and therefore these degrons were not explored further. However, the Hsl1-derived 'super D-box' peptide (D1: Ac-GRAALSDITN-NH_2) provided promising initial results, and further D-box peptides were designed and synthesised to establish a SAR. This iterative design process revealed the optimal natural sequence to be Ac-RAPLSDITN-NH_2 (D19), which bound with an affinity of 5.87 ± 0.11 μM to Cdc20. In the process of determining this high affinity natural amino acid peptide, we uncovered structural determinants that result in improved D-box degron affinity to Cdc20. Examples of this include the P3A mutation resulting in approximately 3-fold improvement of affinity; G5S mutation resulting in around 3-fold improvement in affinity and Isoleucine being the most efficient residue at position 7 for binding, compared to the other two aliphatic amino acids, Valine and Leucine. In addition, the use of tri-methylated (C_3) and tri-fluorinated (F_3) non-natural amino acids in place of the canonical leucine residue resulted in a remarkable improvement in the resulting peptides: D7, D12, D20 and D21. These modifications resulted in binding affinities of 902 ± 14 and 524 ± 14 nM for D20 (Ac-RAPC_3SDITN-NH_2) and D21 (Ac-RAPF_3SDITN-NH2), respectively which are approximately 6-fold and 11-fold improved from the Leucine-containing parental peptide, D19.

It was then demonstrated that the highest affinity D-box peptides: D21, D20, D7 and D19 were able to bind exogenous full-length Cdc20 in a split-luciferase CETSA assay indicating that these peptides are able to engage Cdc20 within the context of the cellular environment. We were also able to demonstrate that the C-box hydrocarbon-stapled ligands, C6e and C6l were potent destabilising peptides against the Apc8 subunit of the core APC/C. This result, in conjunction with the *in vitro* ubiquitination experiments performed by Dr. Thomas Tischer in Prof. David Barford's laboratory demonstrated for the first time that these peptides could also function as potent inhibitors against APC/C^{Cdc20}-mediated substrate ubiquitination. These ligands were significantly more potent than the current alternative, TAME and have the potential to be further developed either rationally or through focused screening. In addition, our lead D-box ligands,

D20 and D21, which were weaker in affinity as determined by SPR, also outperformed the small molecule Apcin in their ability to inhibit Cyclin B1 ubiquitination. Lastly, we successfully crystallised and determined the co-crystal structures of three D-box peptides (D7, D20 and D21) bound to Cdc20 with resolutions of 1.92 Å, 1.46 Å and 1.51 Å, respectively. These data provide evidence of the binding mode and confirmed the SAR and the models established throughout the design process during their development. Notably, these are the first human Cdc20 crystal structures with D-box peptides bound, despite extensive attempts previously in the literature.

Overall, the rational design approach has proved to be able to develop potent D-box and C-box peptide inhibitors, which target the Anaphase Promoting Complex/Cyclosome. These first in-class peptide inhibitory molecules have the potential to be further explored, modified and developed to answer countless further questions pertaining to the overall efficacy of suitability these molecules may have as therapeutic candidates to eventually treat a wide range of cancers. Until then, these initial molecules may serve a useful role modulating the APC/C^{Cdc20} ubiquitination activity in the context of further understanding APC/C function in the cell.

7.0.3 Future work

Due to the high homology between Cdc20 and Cdh1, it is very likely that the D-box peptides, which were optimised for binding to Cdc20, could also bind to Cdh1. However, in the context of cancers in which Cdc20 is highly over-expressed [163] and Cdh1 expression is down-regulated [253], such off-target effects of D-box peptides may be negligible or even irrelevant for the overall efficacy of the ligand. Although Cdh1 is the obvious secondary target, probing other off-target effects is also worthwhile and necessary to assess possible negative side-effects that may occur as a result. The combination of CETSA with proteomic profiling could be used identify proteins that are stabilised due to the inhibition of the APC/C [254–256]. At present, a joint PhD student between the Itzhaki and Lindon labs is investigating the sequence and structural requirement of the degrons of the APC/C^{Cdh1}, and she will examine whether the peptides presented here can act as Cdh1 inhibitors. She is also exploring whether more potent peptide-based inhibitors can be developed by incorporating multiple degrons in a single protein to function as multivalent biologic inhibitors.

With the advent of the *in vitro* ubiquitination experiments providing evidence of inhibitory actions for peptides D20, D21 and C6e, further functionalisation of these ligands to traverse the cell membrane would be highly worthwhile. If we can achieve this, these molecules could be further characterised to examine their functional output, i.e. whether they can inhibit mitotic exit. In addition, the curious results of the C-box ligands in particular could be further examined using negative stain electron microscopy or even cryogenic electron microscopy to determine whether these ligand binding events cause broader structural re-arrangements of the APC/C. This would be particularly interesting in the case of C6l, which produced the increased levels of Cyclin B1 ubiquitination.

Similar to what was described above for Tankyrase peptides, library display approaches could also be used to generate higher affinity ligands for Cdc20 and simultaneously counter-screen against Cdh1, if Cdc20-specific inhibitors are desired. This could result in more potent ligands that target the D-box, KEN-box and ABBA motif degrons that adhere towards the known consensus sequences, but it could also possibly generate new and diverse ligands and binding conformation for use as tools or further development towards therapeutics. However, library-based methods are unlikely to be successful for the development of peptide inhibitors based on the C-box peptide, given that the Apc8 and Apc3 subunit cannot be produced outside the context of the APC/C. Instead the split-luciferase CETSA approach could potentially be combined with peptide libraries to screen for stabilising or indeed as we observed, destabilising Apc8 and Apc3 ligands.

In these studies, only a limited set of hydrocarbon staples and non-natural amino acids were investigated during the PhD. However, the possibilities of expanding the chemical space are vast and may include alternative chemical constraints and modifications that could lead to more potent ligands, improved efficacy and future drug candidates. Utilising our new crystal structure data, models were generated for a Lanthionine-containing peptide to constrain between position 2 to position 5 7.1. This new chemistry might be more amenable given the longer bond lengths of the C-S bonds compared to C-C bonds, that are adopted in the hydrocarbon-stapled peptides.

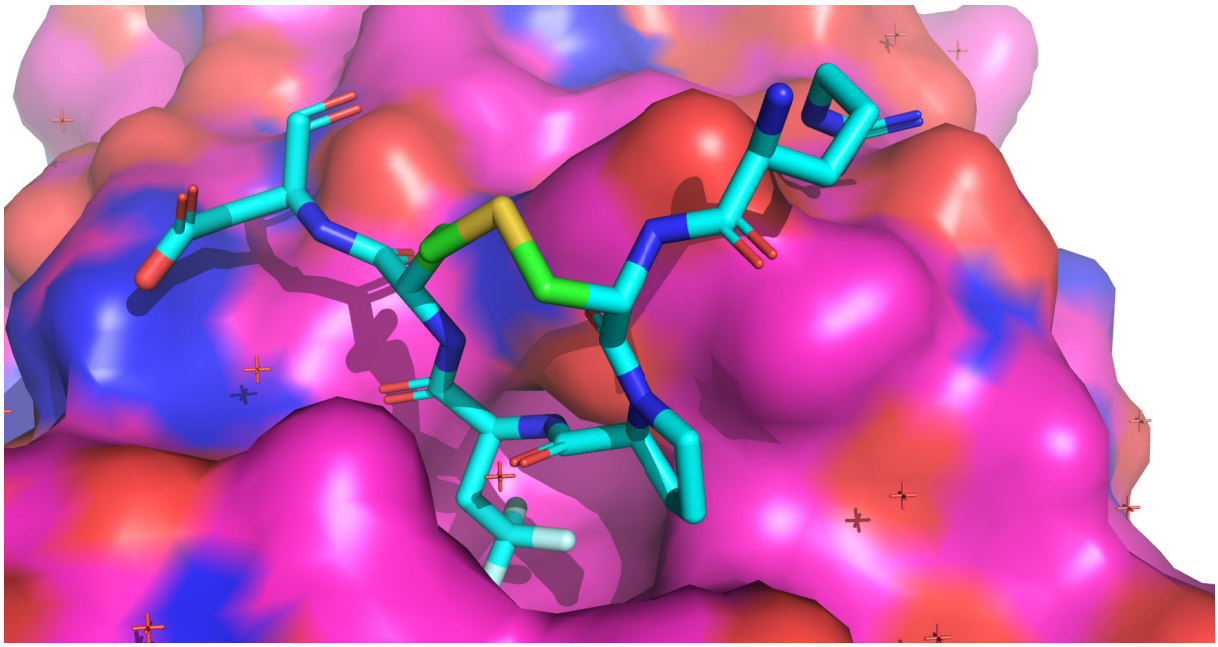


Figure 7.1: Models generated for possible lanthionine based chemistries to explore new chemical space to staple between position 2 and 5 of D-box peptides. Cdc20 in magenta surface, the D21 parental peptide coloured in cyan, stick representation with the new staple modelin green. PDB: Cdc20-D21 structure from this work.

Lastly, the field of targeted protein degradation (TPD) is rapidly evolving and like discussed at the during the introduction, the pharmaceutical industry rapidly became limited in the scope of possible degradation mechanisms. Of the >600 predicted E3 ubiquitin ligases that could be used for TPD, very few are commonly used, such as Cereblon, VHL, IAPs and Mdm2 have been utilised [257]. The APC/C offers a unique position to be utilised for TPD application due to its already wide range of (>60) known substrates that it can degrade. This characteristic could be highly valuable for these applications and the new ligands and knowledge based from these studies could pave the way for utilising this E3 ubiquitin ligase for a variety of pathological diseases.

Appendix A

Protein sequences

Key:
Affinity tags = X
Protease cleavage site = X, | indicates site of peptide bond cleavage.

His_6-TEV protease (S219V)

MGHHHHHHGESLFKGPRDYNPISSTICHLTNESDGHTTSLYGIGFGPFIITNKHLF
RRNNGTLLVQSLHGVFKVKNTTTLQQHLIDGRDMIIIRMPKDFPPFPQKLKFREP
QREERICLVTTNFQTKSMSSMVSDTSCTFPSSDGIFWKHWIQTKDGQCGSPLVST
RDGFIVGIHSASNFTNTNNYFTSVPKNFMELLTNQEAQQWVSGWRLNADSVLW
GGHKVFMVKPEEPFQPVKEATQLMNELVYSQ

His_6-TEV-Mdm2 (6-125)

MRGSHHHHHHNNNNNNNNNNENLYFQ|GGSMSVPTDGAVTTSQIPASEQETLVR
PKPLLLKLLKSVGAQKDTYTMKEVLFYLGQYIMTKRLYDEKQQHIVYCSNDLLG
DLFGVPSFSVKEHRKIYTMIYRNLVVVNQQESSDSGTSVSEN

His_6-thrombin-TNKS1 ARC2-3 (315-662)

MRGSHHHHHHGLVPR|GSGKSALDLADPSAKAVLTGEYKKDELLEAARSGNEEK
LMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLV
PLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHG
ADPTLVNCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALE
IINFKQPQSHETALHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAA
ERAHNDVMEVLHKHGAKMNALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSII
SLQGFTAAQMGNEAVQQILSESTPIRTSDVDYRLLEASKAGD

His_6-thrombin-TNKS2 ARC4 (488-649)

MAHHHHHHSDLSDLVPR|GSGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDI
EGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVA
ELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTP
LDLVKDGDTDIQDLLRGDAAL

GST-thrombin-TNKS2 ARC4 (488-649)

MMSPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFP
NLPYYIDGDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRI
AYSKDFETLKVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVV
LYMDPMCLDAFPKLVCFKKRIEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHP
PKSDLVPR|GSGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLH
FAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAV
VNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDT
DIQDLLRGDAAL

GST-thrombin-TNKS2 ARC4 (488-649, G558D)

MMSPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFP
NLPYYIDGDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRI
AYSKDFETLKVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVV
LYMDPMCLDAFPKLVCFKKRIEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHP
PKSDLVPR|GSGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLH
FAAGYNRVSVVEYLLQHGADVHAKDKDGLVPLHNACSYGHYEVAELLVKHGAV
VNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDT
DIQDLLRGDAAL

GST-thrombin-TNKS2 ARC4 (488-649, G558S)

MMSPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFP
NLPYYIDGDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRI
AYSKDFETLKVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVV
LYMDPMCLDAFPKLVCFKKRIEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHP
PKSDLVPR|GSGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLH
FAAGYNRVSVVEYLLQHGADVHAKDKSGLVPLHNACSYGHYEVAELLVKHGAV
VNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDT
DIQDLLRGDAAL

GST-thrombin-TNKS2 ARC4 (488-649, G558T)

MMSPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFP
NLPYYIDGDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRI
AYSKDFETLKVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVV
LYMDPMCLDAFPKLVCFKKRIEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHP
PKSDLVPR|GSGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLH
FAAGYNRVSVVEYLLQHGADVHAKDKTGLVPLHNACSYGHYEVAELLVKHGAV
VNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDT
DIQDLLRGDAAL

His$_6$-MBP-TEV-Cdc20 (161-477)

MSYYHHHHHHDYDIPTTKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTV
EHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYP
FTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSA
LMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIK
NKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQ
PSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYE
EELAKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEALK
DAQTNSSSNNNNNNNNNNENLYFQ|GASRKTCRYIPSLPDRILDAPEIRNDYYLNL
VDWSSGNVLAVALDNSVYLWSASSGDILQLLQMEQPGEYISSVAWIKEGNYLAVG
TSSAEVQLWDVQQQKRLRNMTSHSARVGSLSWNSYILSSGSRSGHIHHHDVRVAE
HHVATLSGHSQEVCGLRWAPDGRHLASGGNDNLVNVWPSAPGEGGWVPLQTF
TQHQGAVKAVAWCPWQSNVLATGGGTSDRHIRIWNVCSGACLSAVDAHSQVCS
ILWSPHYKELISGHGFAQNQLVIWKYPTMAKVAELKGHTSRVLSLTMSPDGATV
ASAAADETLRLWRCFELD

His$_6$-MBP-TEV-Cdc20 (165-477)

MSYYHHHHHHDYDIPTTKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTV
EHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYP
FTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSA
LMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIK
NKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQ
PSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYE
EELAKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEALK
DAQTNSSSNNNNNNNNNNENLYFQ|GACRYIPSLPDRILDAPEIRNDYYLNLVDWS
SGNVLAVALDNSVYLWSASSGDILQLLQMEQPGEYISSVAWIKEGNYLAVGTSSA
EVQLWDVQQQKRLRNMTSHSARVGSLSWNSYILSSGSRSGHIHHHDVRVAEHHV

ATLSGHSQEVCGLRWAPDGRHLASGGNDNLVNVWPSAPGEGGWVPLQTFTQH
QGAVKAVAWCPWQSNVLATGGGTSDRHIRIWNVCSGACLSAVDAHSQVCSILWS
PHYKELISGHGFAQNQLVIWKYPTMAKVAELKGHTSRVLSLTMSPDGATVASAA
ADETLRLWRCFELD

His_6-TEV-Cdc20 (161-477)

MSYYHHHHHHDYDIPTTNSSSNNNNNNNNNNENLYFQ|GASRKTCRYIPSLPDRIL
DAPEIRNDYYLNLVDWSSGNVLAVALDNSVYLWSASSGDILQLLQMEQPGEYISS
VAWIKEGNYLAVGTSSAEVQLWDVQQQKRLRNMTSHSARVGSLSWNSYILSSGS
RSGHIHHHDVRVAEHHVATLSGHSQEVCGLRWAPDGRHLASGGNDNLVNVWPS
APGEGGWVPLQTFTQHQGAVKAVAWCPWQSNVLATGGGTSDRHIRIWNVCSG
ACLSAVDAHSQVCSILWSPHYKELISGHGFAQNQLVIWKYPTMAKVAELKGHTS
RVLSLTMSPDGATVASAAADETLRLWRCFELD

His_6-TEV-Cdc20 (165-477)

MSYYHHHHHHDYDIPTTNSSSNNNNNNNNNNENLYFQ|GACRYIPSLPDRILDAPE
IRNDYYLNLVDWSSGNVLAVALDNSVYLWSASSGDILQLLQMEQPGEYISSVAWI
KEGNYLAVGTSSAEVQLWDVQQQKRLRNMTSHSARVGSLSWNSYILSSGSRSGH
IHHHDVRVAEHHVATLSGHSQEVCGLRWAPDGRHLASGGNDNLVNVWPSAPGE
GGWVPLQTFTQHQGAVKAVAWCPWQSNVLATGGGTSDRHIRIWNVCSGACLS
AVDAHSQVCSILWSPHYKELISGHGFAQNQLVIWKYPTMAKVAELKGHTSRVLSL
TMSPDGATVASAAADETLRLWRCFELD

HiBiT-Cdc20 (1-499)

MGSVSGWRLFKKISGSMAQFAFESDLHSLLQLDAPIPNAPPARWQRKAKEAAGP
APSPMRAANRSHSAGRTPGRTPGKSSSKVQTTPSKPGGDRYIPHRSAAQMEVAS
FLLSKENQPENSQTPTKKEHQKAWALNLNGFDVEEAKILRLSGKPQNAPEGYQN
RLKVLYSQKATPGSSRKTCRYIPSLPDRILDAPEIRNDYYLNLVDWSSGNVLAVA
LDNSVYLWSASSGDILQLLQMEQPGEYISSVAWIKEGNYLAVGTSSAEVQLWDVQ
QQKRLRNMTSHSARVGSLSWNSYILSSGSRSGHIHHHDVRVAEHHVATLSGHSQE
VCGLRWAPDGRHLASGGNDNLVNVWPSAPGEGGWVPLQTFTQHQGAVKAVA
WCPWQSNVLATGGGTSDRHIRIWNVCSGACLSAVDAHSQVCSILWSPHYKELISG
HGFAQNQLVIWKYPTMAKVAELKGHTSRVLSLTMSPDGATVASAAADETLRLW
RCFELDPARRREREKASAAKSSLIHQGIR

Cdc20 (1-499)-HiBiT

MGSMAQFAFESDLHSLLQLDAPIPNAPPARWQRKAKEAAGPAPSPMRAANRSHS
AGRTPGRTPGKSSSKVQTTPSKPGGDRYIPHRSAAQMEVASFLLSKENQPENSQT
PTKKEHQKAWALNLNGFDVEEAKILRLSGKPQNAPEGYQNRLKVLYSQKATPG
SSRKTCRYIPSLPDRILDAPEIRNDYYLNLVDWSSGNVLAVALDNSVYLWSASSGD
ILQLLQMEQPGEYISSVAWIKEGNYLAVGTSSAEVQLWDVQQQKRLRNMTSHSA
RVGSLSWNSYILSSGSRSGHIHHHDVRVAEHHVATLSGHSQEVCGLRWAPDGRHL
ASGGNDNLVNVWPSAPGEGGWVPLQTFTQHQGAVKAVAWCPWQSNVLATGG
GTSDRHIRIWNVCSGACLSAVDAHSQVCSILWSPHYKELISGHGFAQNQLVIWKY
PTMAKVAELKGHTSRVLSLTMSPDGATVASAAADETLRLWRCFELDPARRRER
EKASAAKSSLIHQGIRGSVSGWRLFKKISGS

HiBiT-Apc3 (1-824)

MGSVSGWRLFKKISGSMTVLQEPVQAAIWQALNHYAYRDAVFLAERLYAEVHSE
EALFLLATCYYRSGKAYKAYRLLKGHSCTTPQCKYLLAKCCVDLSKLAEGEQILS
GGVFNKQKSHDDIVTEFGDSACFTLSLLGHVYCKTDRLAKGSECYQKSLSLNPFL
WSPFESLCEIGEKPDPDQTFKFTSLQNFSNCLPNSCTTQVPNHSLSHRQPETVLTE
TPQDTIELNRLNLESSNSKYSLNTDSSVSYIDSAVISPDTVPLGTGTSILSKQVQNKP
KTGRSLLGGPAALSPLTPSFGILPLETPSPGDGSYLQNYTNTPPVIDVPSTGAPSK
KSVARIGQTGTKSVFSQSGNSREVTPILAQTQSSGPQTSTTPQVLSPTITSPPNALP
RRSSRLFTSDSSTTKENSKKLKMKFPPEIPNRKTKSKTNKGGITQPNINDSLEITK
LDSSIISEGKISTITPQIQAFNLQKAAAEGLMSLLREMGKGYLALCSYNCKEAINILS
HLPSHHYNTGWVLCQIGRAYFELSEYMQAERIFSEVRRIENYRVEGMEIYSTTLW
HLQKDVALSVLSKDLTDMDKNSPEAWCAAGNCFSLQREHDIAIKFFQRAIQVDP
NYAYAYTLLGHEFVLTEELDKALACFRNAIRVNPRHYNAWYGLGMIYYKQEKFS
LAEMHFQKALDINPQSSVLLCHIGVVQHALKKSEKALDTLNKAIVIDPKNPLCKFH
RASVLFANEKYKSALQELEELKQIVPKESLVYFLIGKVYKKLGQTHLALMNFSWA
MDLDPKGANNQIKEAIDKRYLPDDEEPITQEEQIMGTDESQESSMTDADDTQLH
AAESDEF

Apc3 (1-824)-HiBiT

MTVLQEPVQAAIWQALNHYAYRDAVFLAERLYAEVHSEEALFLLATCYYRSGK
AYKAYRLLKGHSCTTPQCKYLLAKCCVDLSKLAEGEQILSGGVFNKQKSHDDIV
TEFGDSACFTLSLLGHVYCKTDRLAKGSECYQKSLSLNPFLWSPFESLCEIGEKP
DPDQTFKFTSLQNFSNCLPNSCTTQVPNHSLSHRQPETVLTETPQDTIELNRLNL
ESSNSKYSLNTDSSVSYIDSAVISPDTVPLGTGTSILSKQVQNKPKTGRSLLGGPA
ALSPLTPSFGILPLETPSPGDGSYLQNYTNTPPVIDVPSTGAPSKKSVARIGQTGT

KSVFSQSGNSREVTPILAQTQSSGPQTSTTPQVLSPTITSPPNALPRRSSRLFTSDS
STTKENSKKLKMKFPPEIPNRKTKSKTNKGGITQPNINDSLEITKLDSSIISEGKIS
TITPQIQAFNLQKAAAEGLMSLLREMGKGYLALCSYNCKEAINILSHLPSHHYNT
GWVLCQIGRAYFELSEYMQAERIFSEVRRIENYRVEGMEIYSTTLWHLQKDVAL
SVLSKDLTDMDKNSPEAWCAAGNCFSLQREHDIAIKFFQRAIQVDPNYAYAYTL
LGHEFVLTEELDKALACFRNAIRVNPRHYNAWYGLGMIYYKQEKFSLAEMHFQ
KALDINPQSSVLLCHIGVVQHALKKSEKALDTLNKAIVIDPKNPLCKFHRASVLFA
NEKYKSALQELEELKQIVPKESLVYFLIGKVYKKLGQTHLALMNFSWAMDLDPK
GANNQIKEAIDKRYLPDDEEPITQEEQIMGTDESQESSMTDADDTQLHAAESDE
FGSVSGWRLFKKISGS

HiBiT-Apc8 (1-591)

MGSVSGWRLFKKISGSMVPVAVTAAVAPVLSINSDFSDLREIKKQLLLIAGLTRER
GLLHSSKWSAELAFSLPALPLAELQPPPPITEEDAQDMDAYTLAKAYFDVKEYDR
AAHFLHGCNSKKAYFLYMYSRYLSGEKKKDDETVDSLGPLEKGQVKNEALRELR
VELSKKHQARELDGFGLYLYGVVLRKLDLVKEAIDVFVEATHVLPLHWGAWLE
LCNLITDKEMLKFLSLPDTWMKEFFLAHIYTELQLIEEALQKYQNLIDVGFSKSSY
IVSQIAVAYHNIRDIDKALSIFNELRKQDPYRIENMDTFSNLLYVRSMKSELSYLAH
NLCEIDKYRVETCCVIGNYYSLRSQHEKAALYFQRALKLNPRYLGAWTLMGHEY
MEMKNTSAAIQAYRHAIEVNKRDYRAWYGLGQTYEILKMPFYCLYYYRRAHQL
RPNDSRMLVALGECYEKLNQLVEAKKCYWRAYAVGDVEKMALVKLAKLHEQL
TESEQAAQCYIKYIQDIYSCGEIVEHLEESTAFRYLAQYYFKCKLWDEASTCAQK
CCAFNDTREEGKALLRQILQLRNQGETPTTEVPAPFFLPASLSANNTPTRRVSPL
NLSSVTP

Apc8 (1-591)-HiBiT

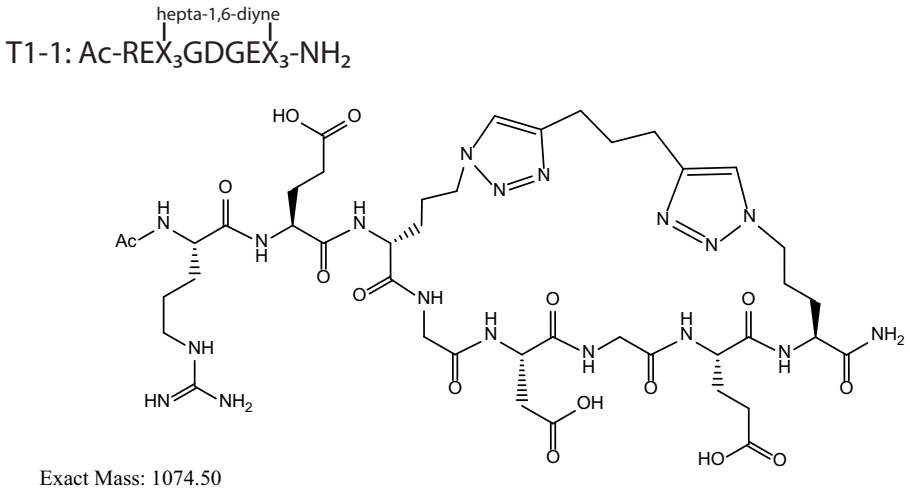
MVPVAVTAAVAPVLSINSDFSDLREIKKQLLLLIAGLTRERGLLHSSKWSAELAFSL
PALPLAELQPPPPITEEDAQDMDAYTLAKAYFDVKEYDRAAHFLHGCNSKKAYF
LYMYSRYLSGEKKKDDETVDSLGPLEKGQVKNEALRELRVELSKKHQARELDGF
GLYLYGVVLRKLDLVKEAIDVFVEATHVLPLHWGAWLELCNLITDKEMLKFLSL
PDTWMKEFFLAHIYTELQLIEEALQKYQNLIDVGFSKSSYIVSQIAVAYHNIRDIDK
ALSIFNELRKQDPYRIENMDTFSNLLYVRSMKSELSYLAHNLCEIDKYRVETCCVI
GNYYSLRSQHEKAALYFQRALKLNPRYLGAWTLMGHEYMEMKNTSAAIQAYRH
AIEVNKRDYRAWYGLGQTYEILKMPFYCLYYYRRAHQLRPNDSRMLVALGECY
EKLNQLVEAKKCYWRAYAVGDVEKMALVKLAKLHEQLTESEQAAQCYIKYIQDI
YSCGEIVEHLEESTAFRYLAQYYFKCKLWDEASTCAQKCCAFNDTREEGKALLR
QILQLRNQGETPTTEVPAPFFLPASLSANNTPTRRVSPLNLSSVTPGSVSGWRLFK
KISGS

Appendix B

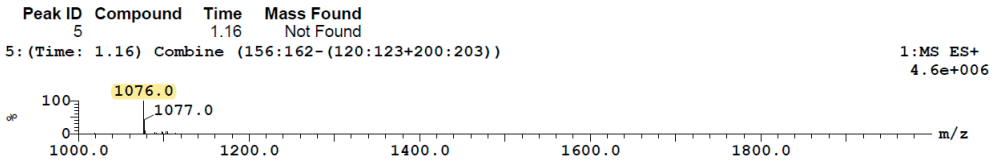
Peptide data

For a subset of peptides, only paper copies of LCMS traces were available and have been scanned in. Some of these data contain handwritten annotations. For all digitial LCMS data, the mass peak corresponding to the M+H, M-H, M+2H and M-2H modes m/z ratios are highlighted. Where it has been observed, both the positive and negative ion modes corresponding to the peptide mass are shown.

T1-1



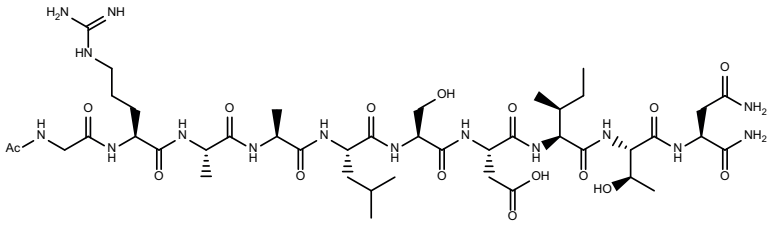
Exact Mass: 1074.50
Molecular Weight: 1075.11



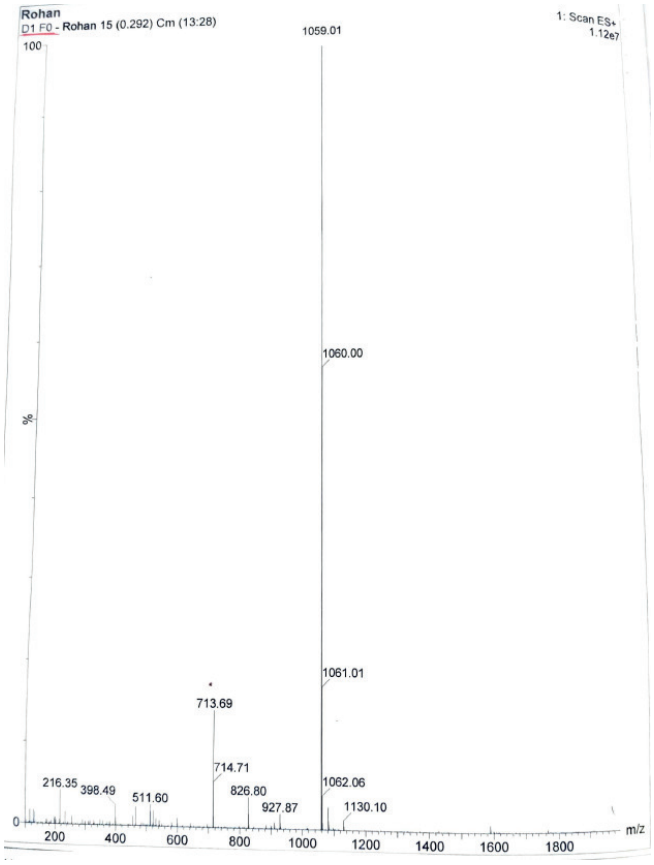
Final analytical HPLC trace was not obtained and the sample no longer exists to generate a chromatogram.

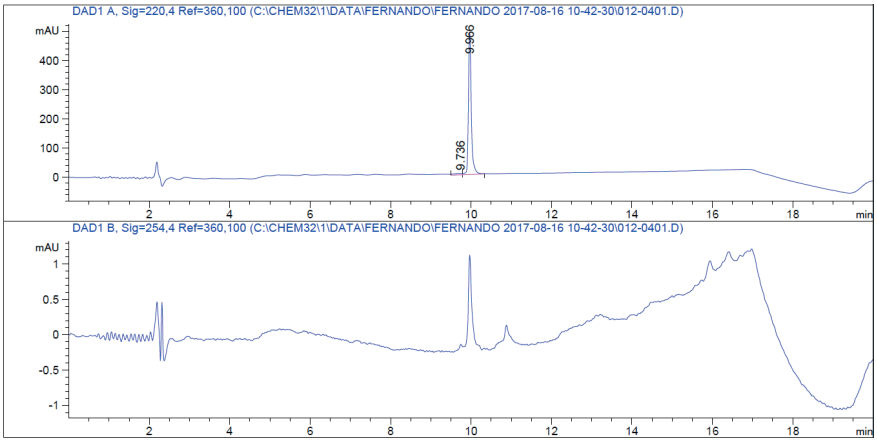
D1

D1: Ac-GRAALSDITN-NH_2



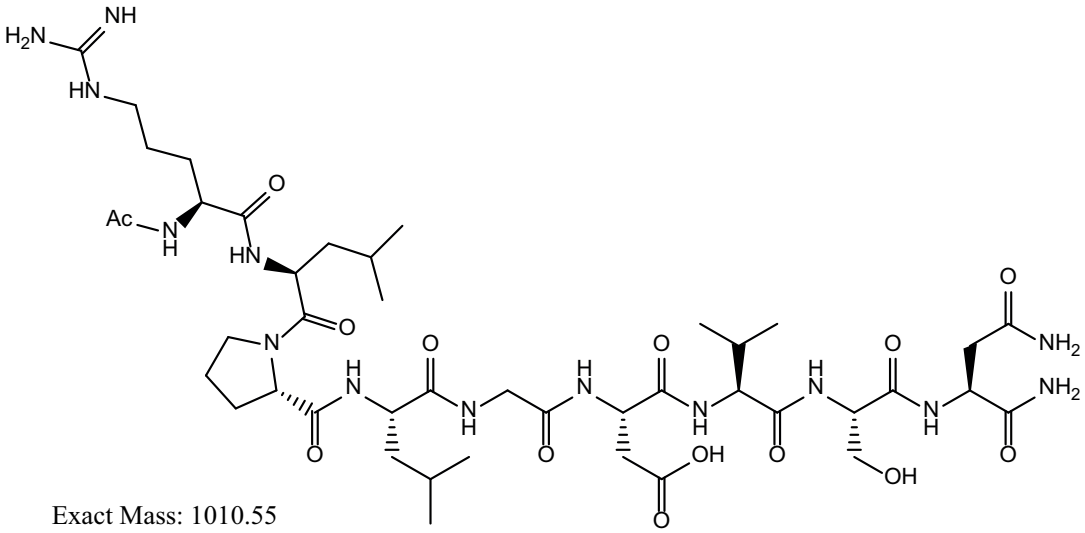
Exact Mass: 1057.55
Molecular Weight: 1058.15





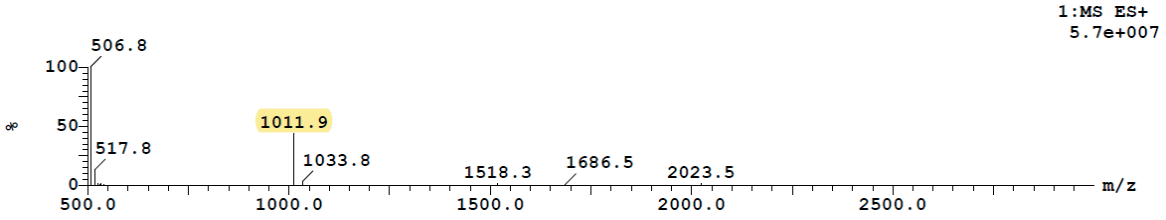
D2

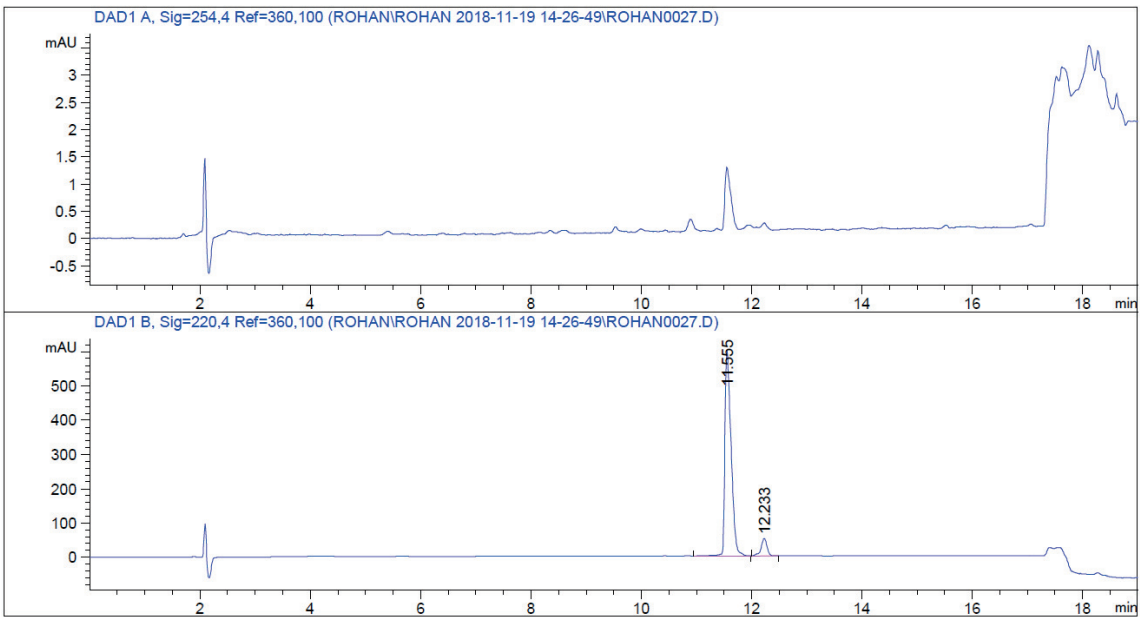
D2: Ac-RLPLGDVSN-NH_2



Exact Mass: 1010.55
Molecular Weight: 1011.13

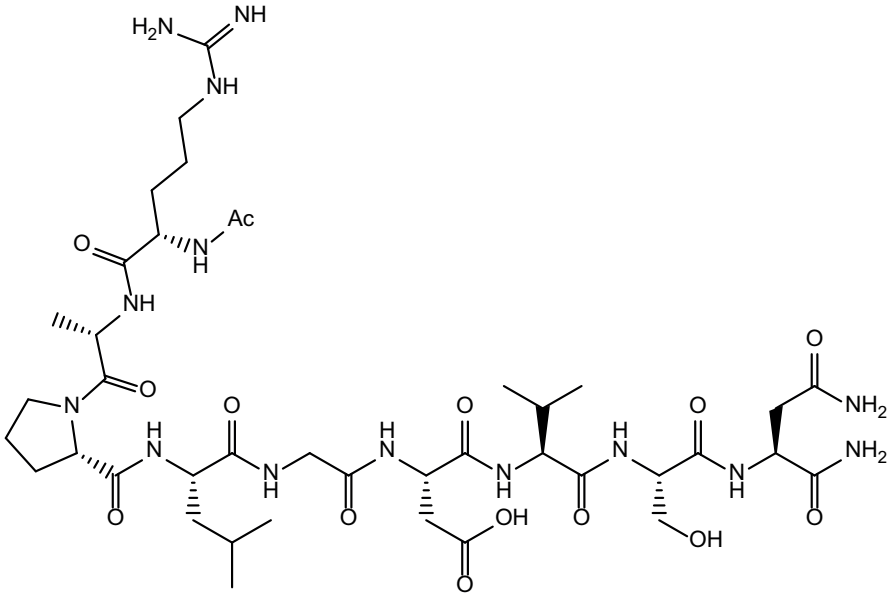
Peak ID	Compound	Time	Mass Found	BPM
10		1.24		507



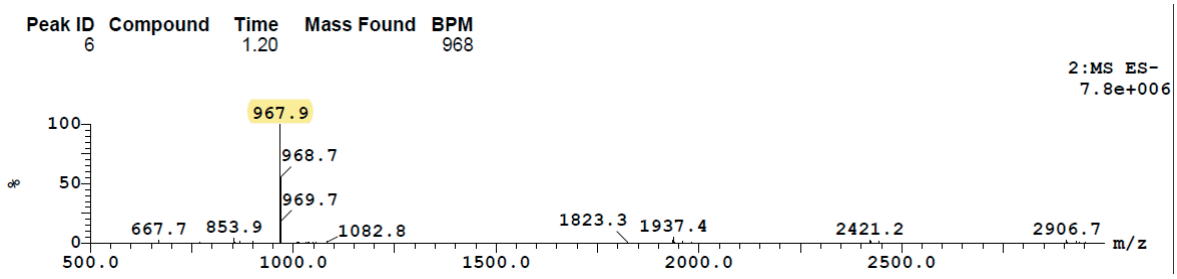


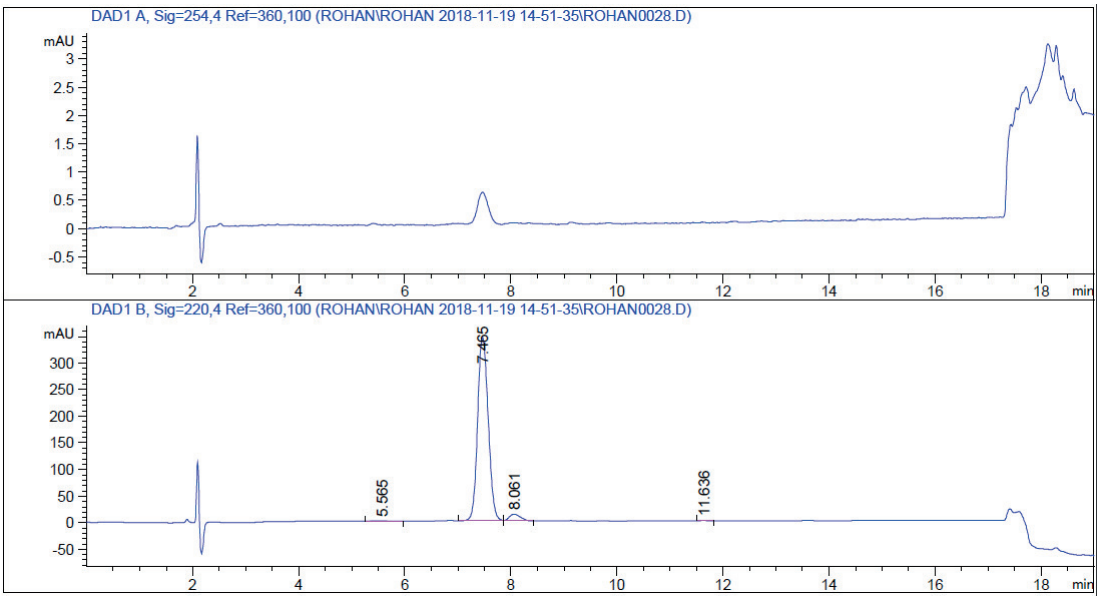
D3

D3: Ac-RAPLGDVSN-NH_2



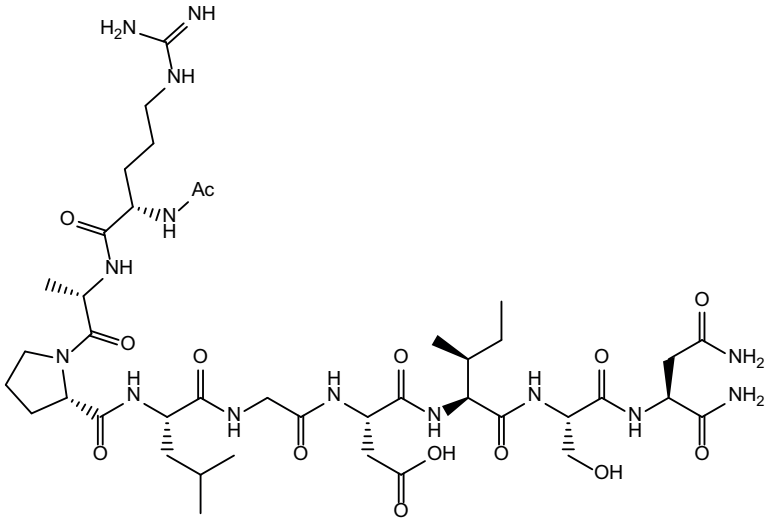
Exact Mass: 968.50
Molecular Weight: 969.05



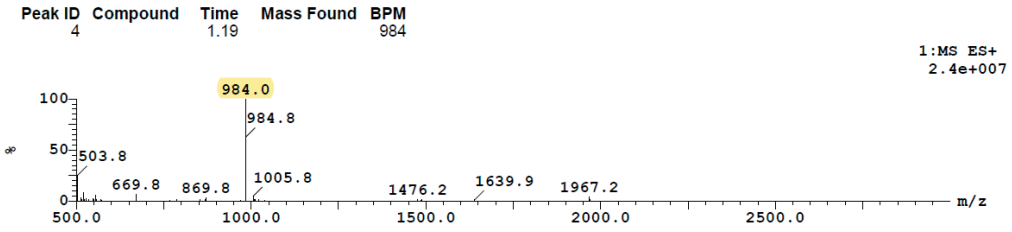


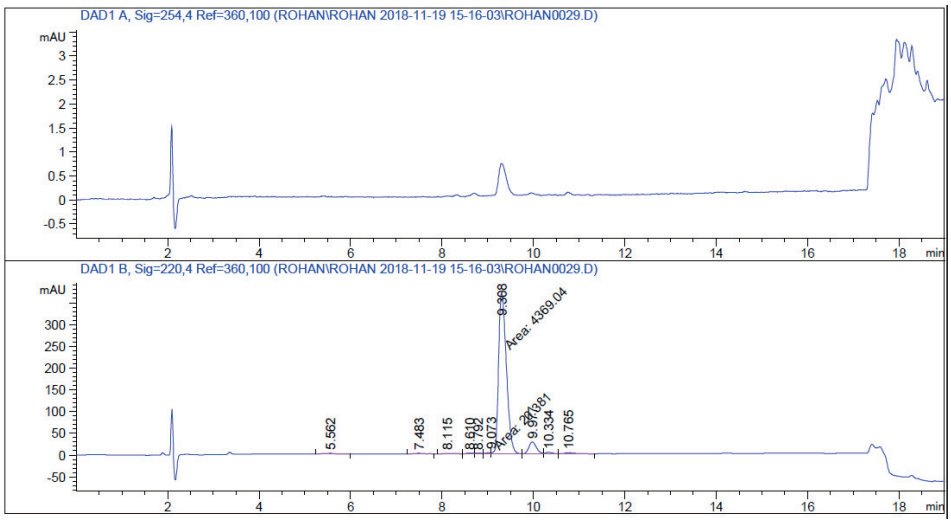
D4

D4: Ac-RAPLGDISN-NH_2



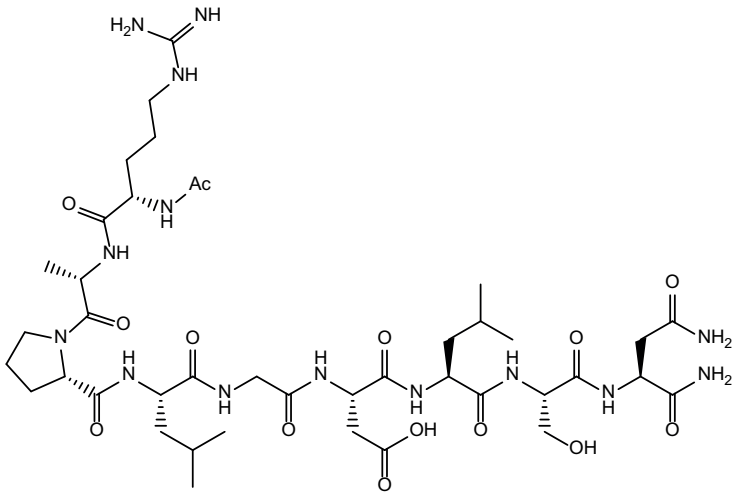
Exact Mass: 982.52
Molecular Weight: 983.08



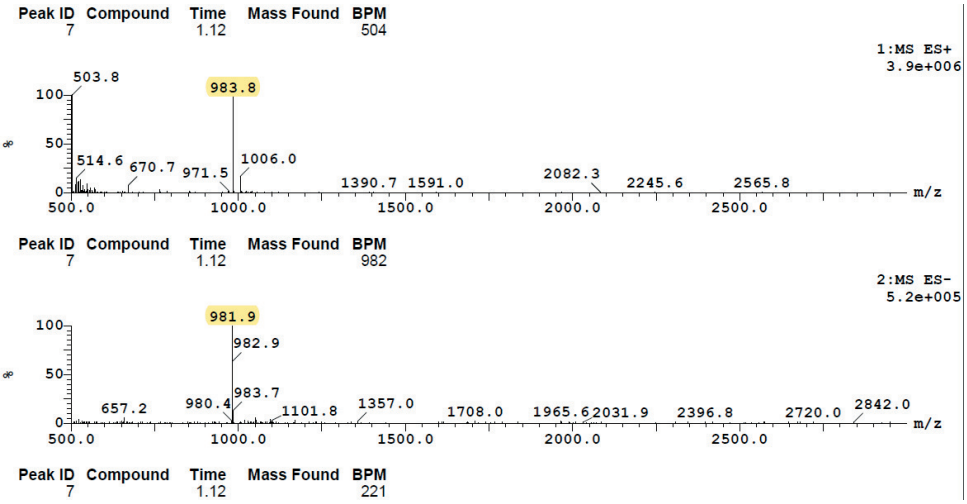


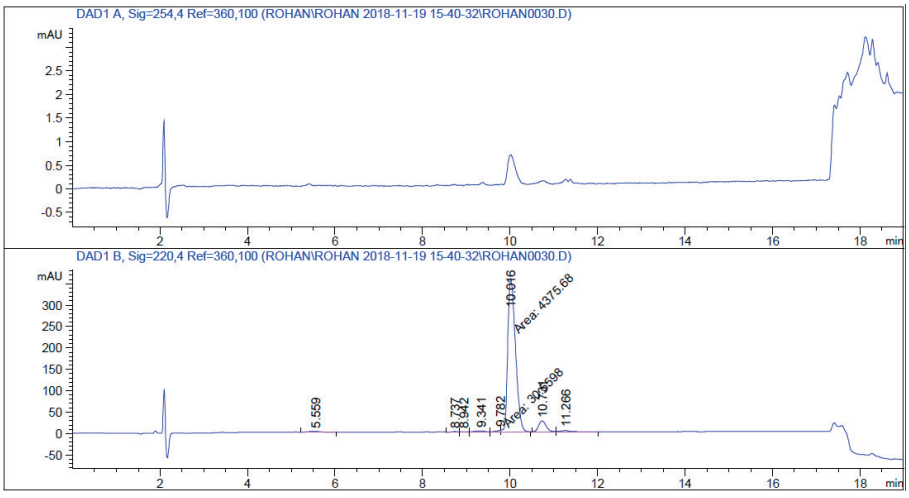
D5

D5: Ac-RAPLGDLSN-NH_2



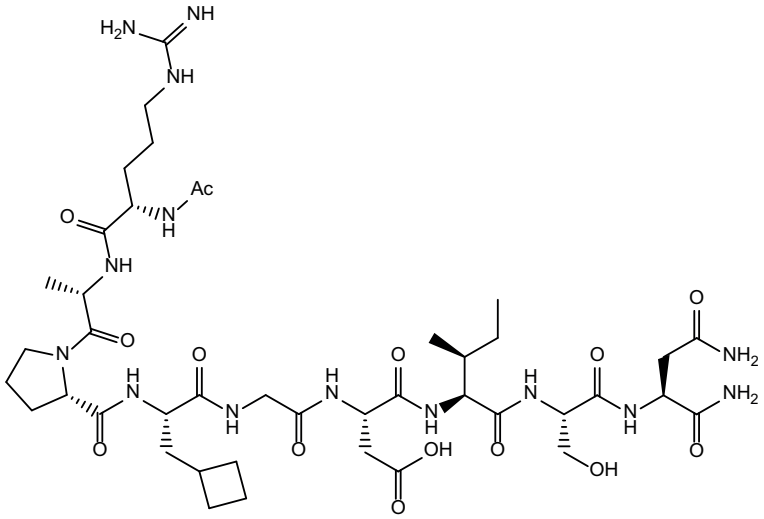
Exact Mass: 982.52
Molecular Weight: 983.08



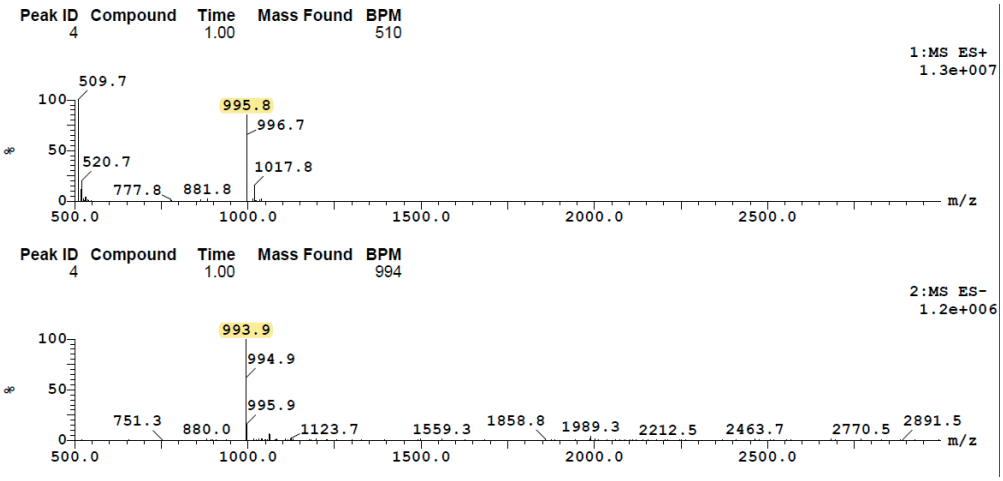


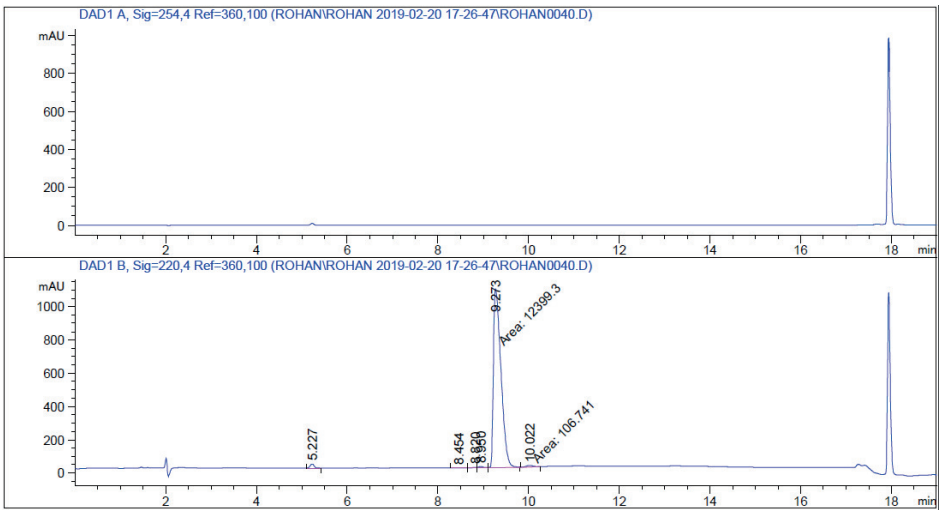
D6

D6: Ac-RAPL**cBuAla**GDISN-NH_2



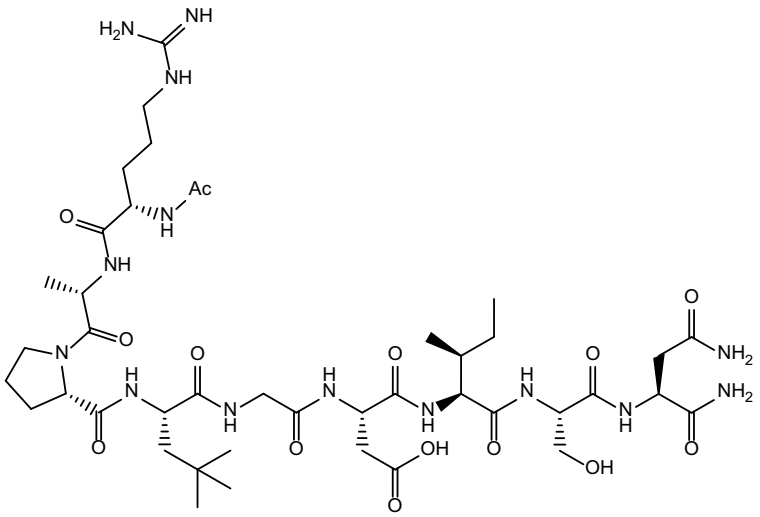
Exact Mass: 994.52
Molecular Weight: 995.09



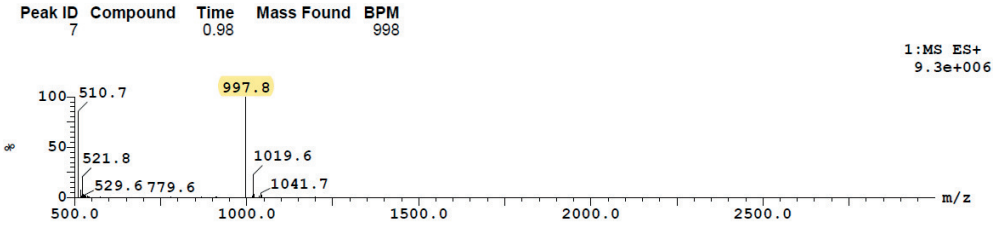


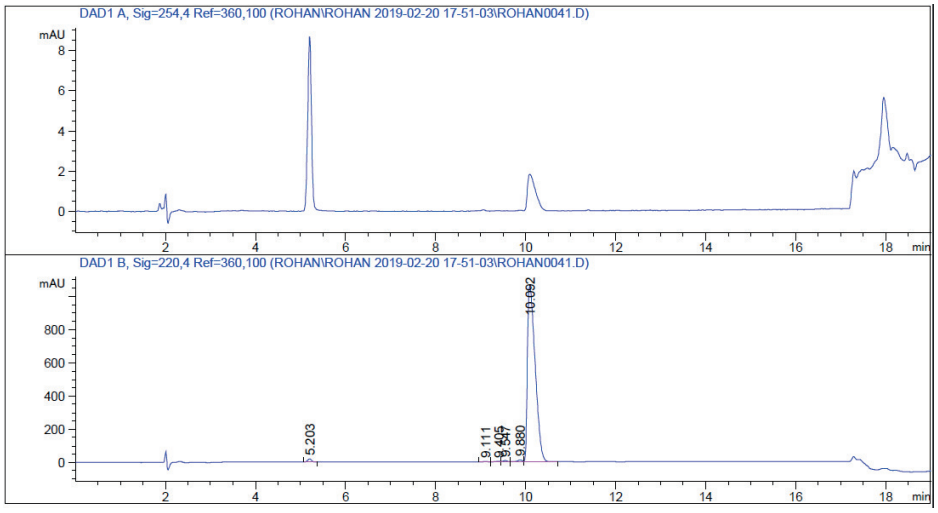
D7

D7: Ac-RAPC_3GDISN-NH_2



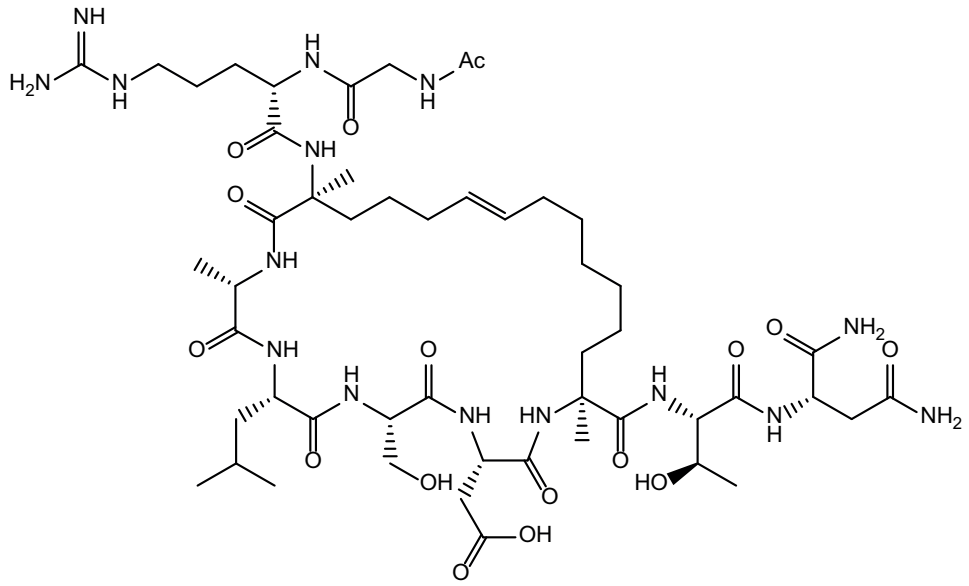
Exact Mass: 996.54
Molecular Weight: 997.11



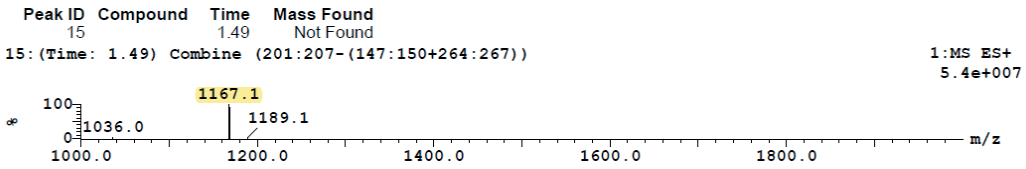


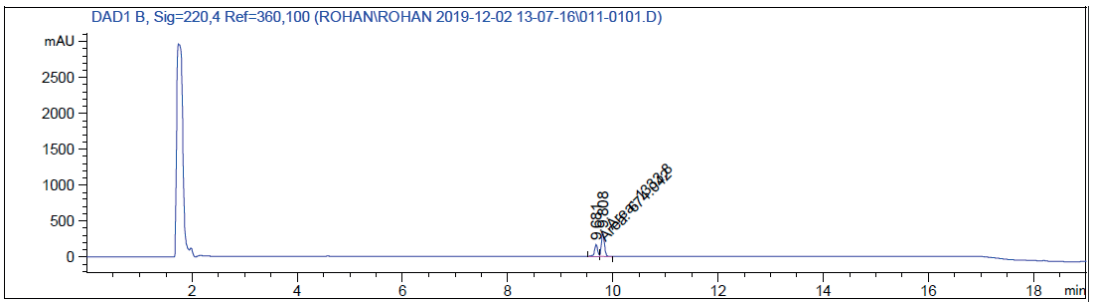
D9

D9: Ac-GR**S$_5$**ALSD**R$_8$**TN-NH$_2$



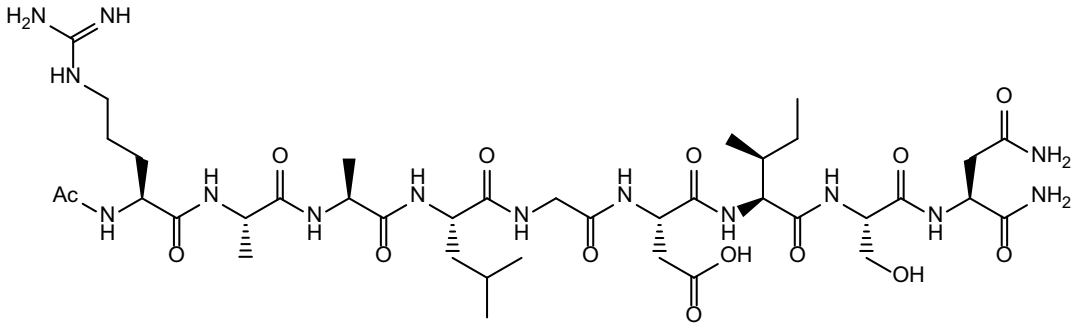
Exact Mass: 1165.65
Molecular Weight: 1166.33



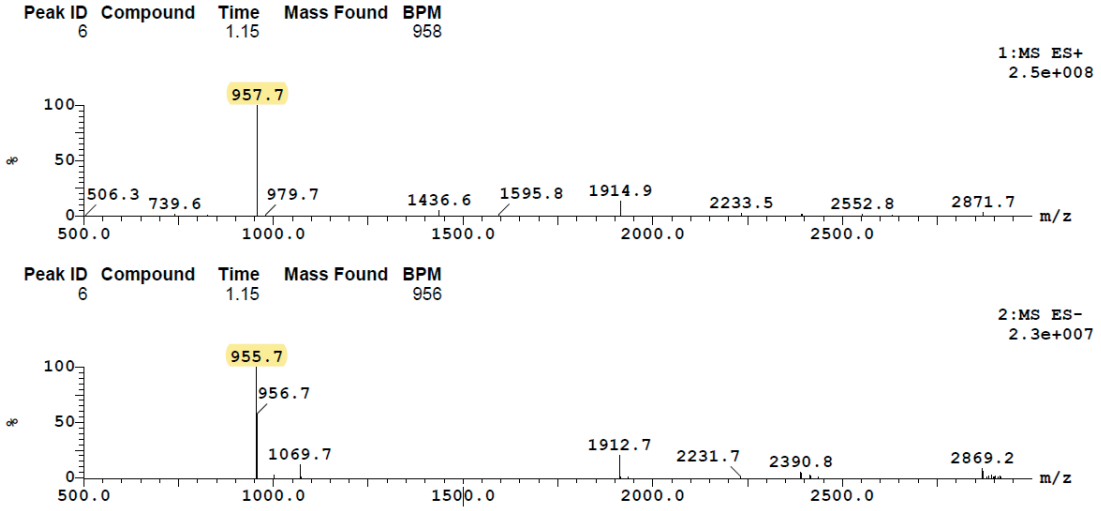


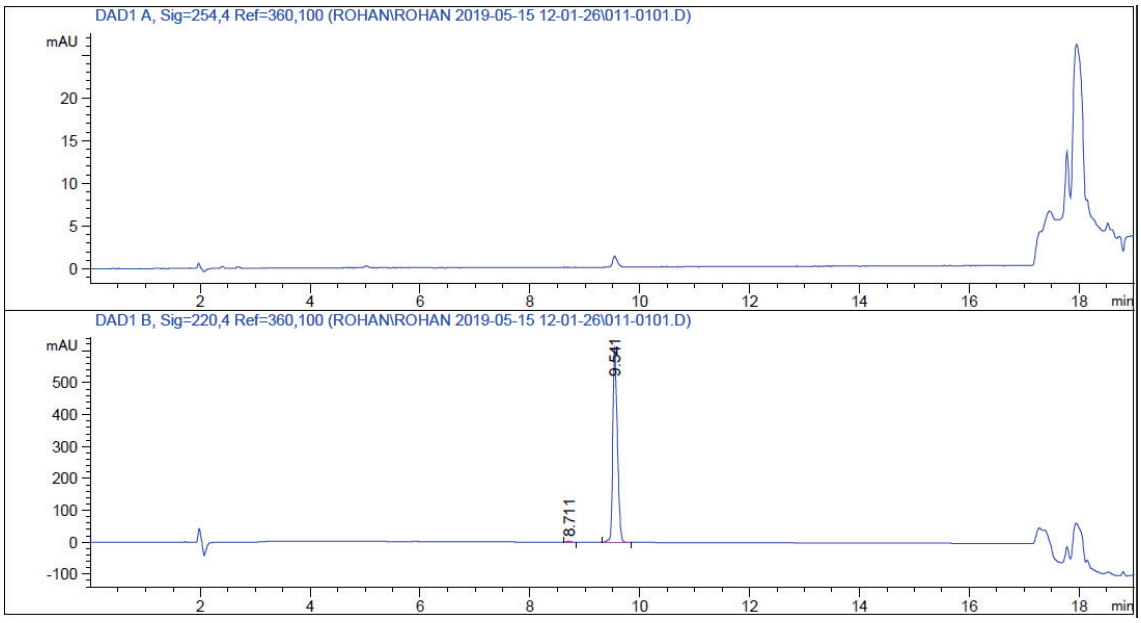
D10

D10: Ac-RAALGDISN-NH_2



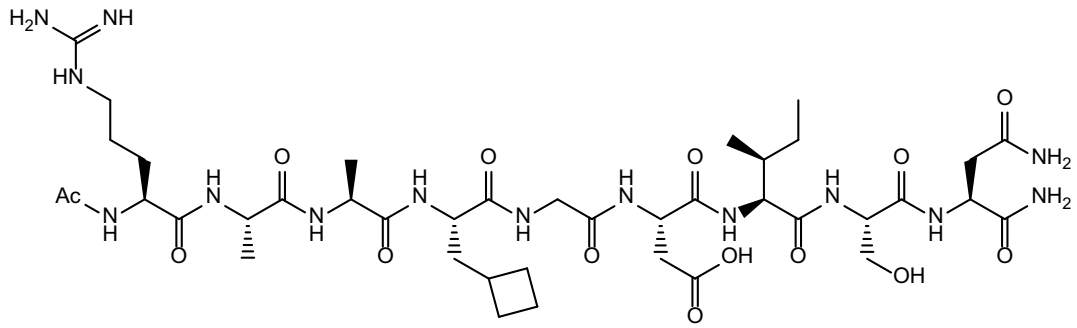
Exact Mass: 956.50
Molecular Weight: 957.04





D11

D11: Ac-RAA**cBuAla**GDISN-NH_2

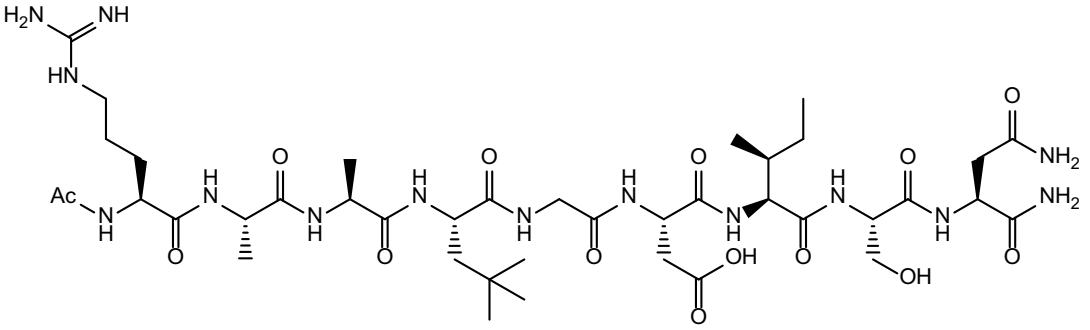


Exact Mass: 968.50
Molecular Weight: 969.05

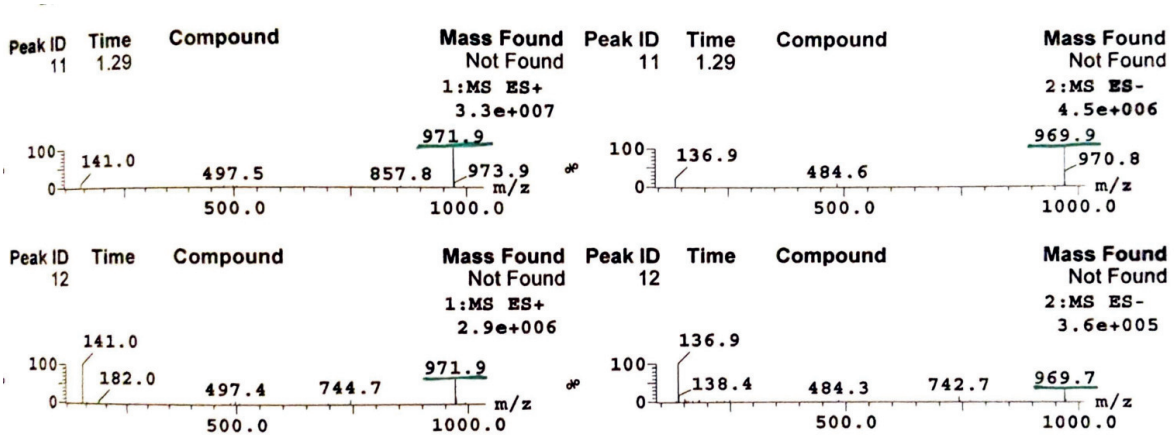
Peptide was insoluble H_2O/MeCN, therefore no characterisation was done with the peptide past its synthesis.

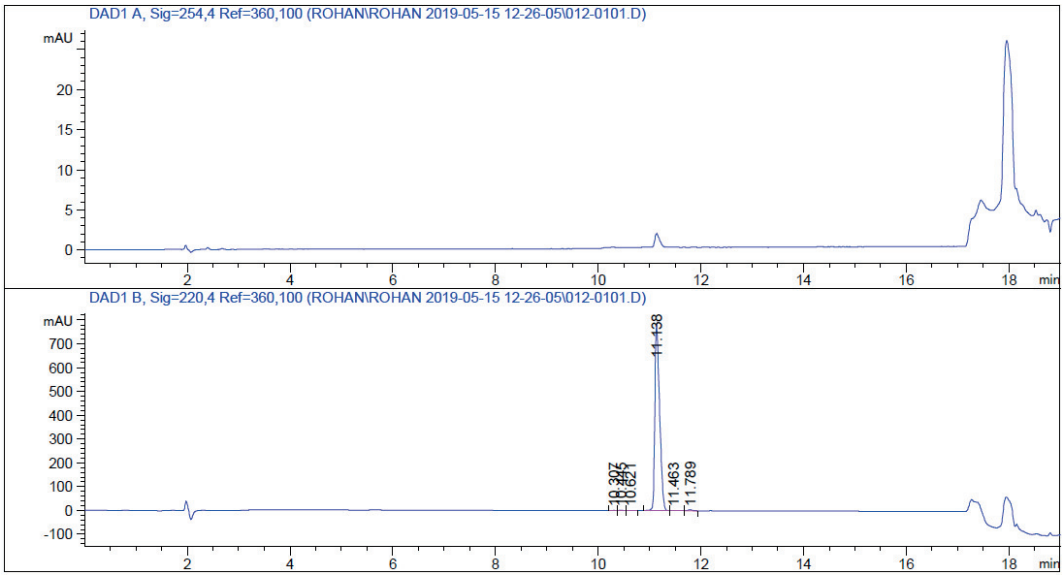
D12

D12: Ac-RAAC_3GDISN-NH_2



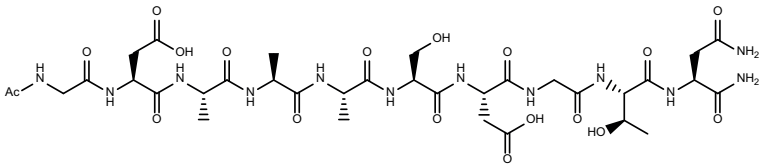
Exact Mass: 970.52
Molecular Weight: 971.07



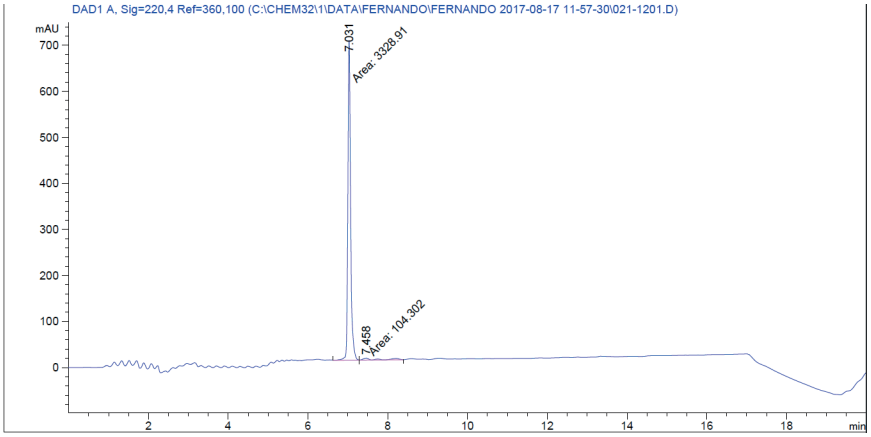


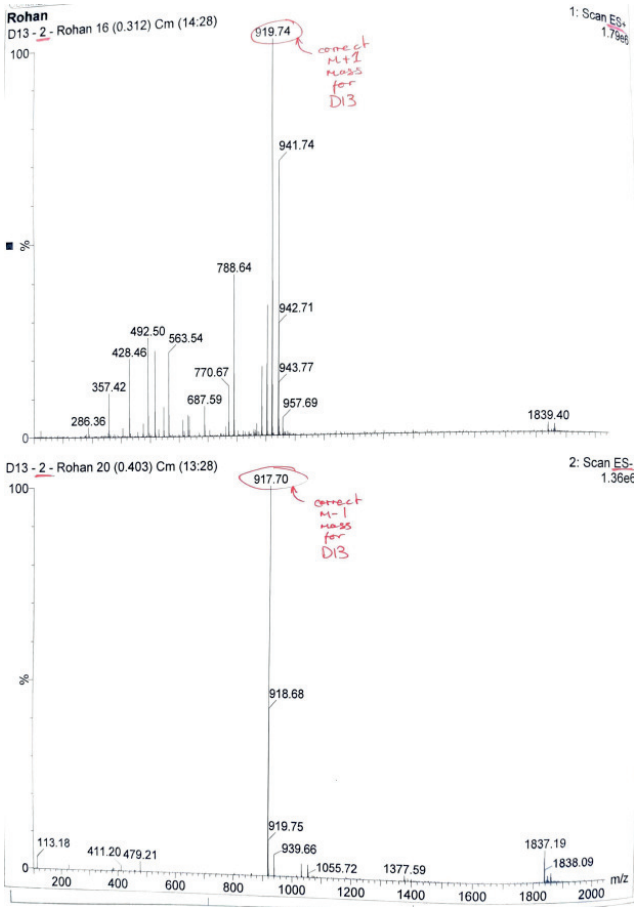
D13

D13: Ac-GDAAASDGTN-NH_2



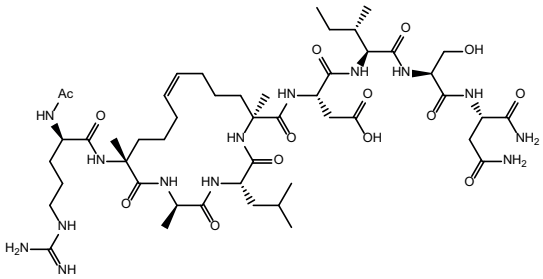
Exact Mass: 918.37
Molecular Weight: 918.86





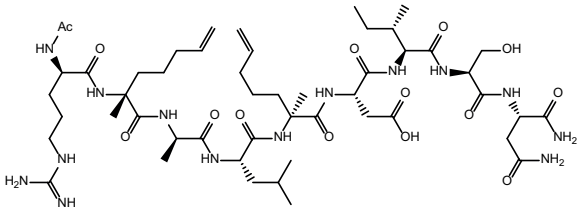
D16

D16: Ac-RS_5ALS_5DISN-NH_2

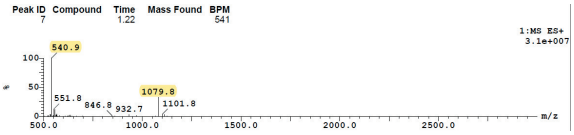


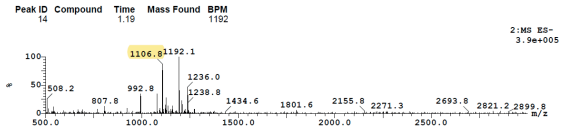
Exact Mass: 1078.61
Molecular Weight: 1079.25

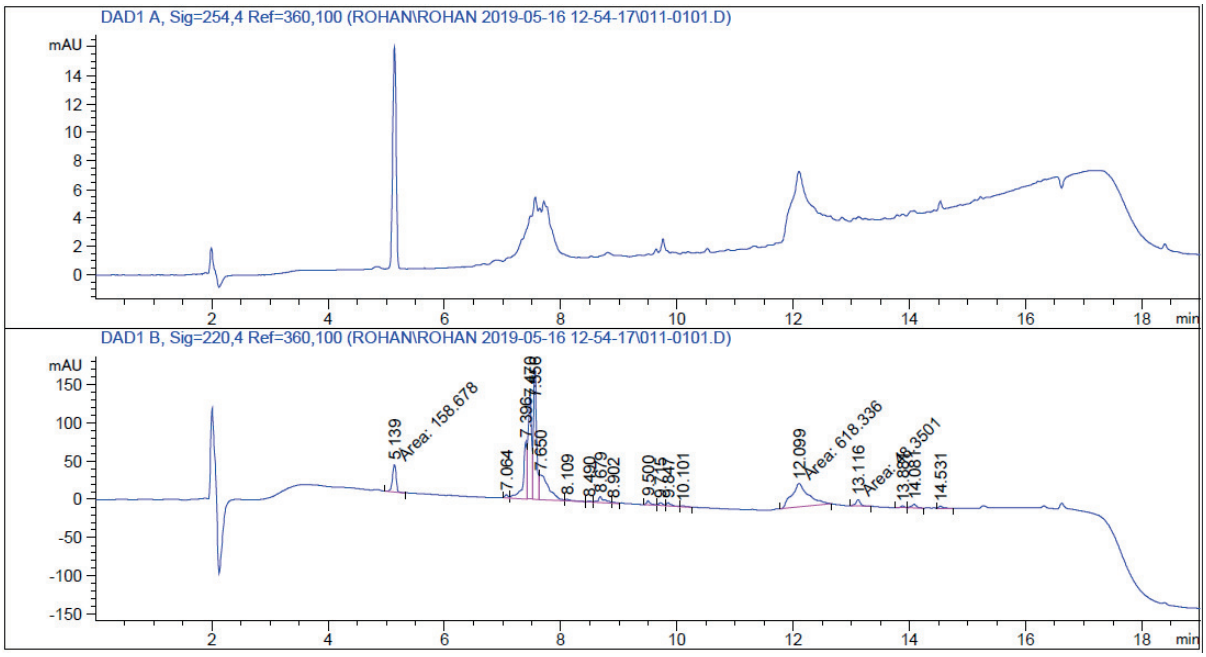
D16 unstapled: Ac-RS_5ALS_5DISN-NH_2



Exact Mass: 1106.64
Molecular Weight: 1107.30

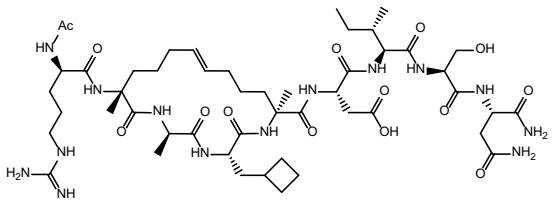






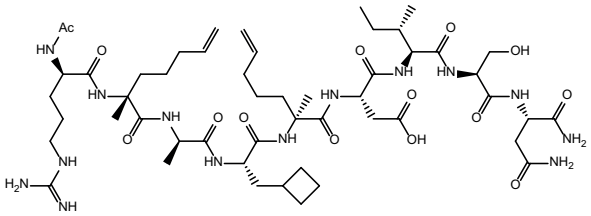
D17

D17: Ac-R**S$_5$**A**cBuAlaS$_5$**DISN-NH$_2$

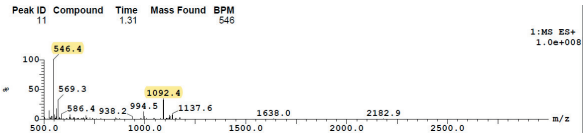


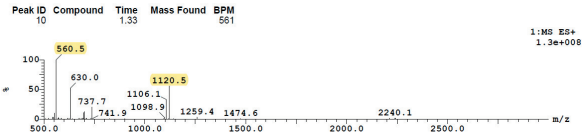
Exact Mass: 1090.61
Molecular Weight: 1091.26

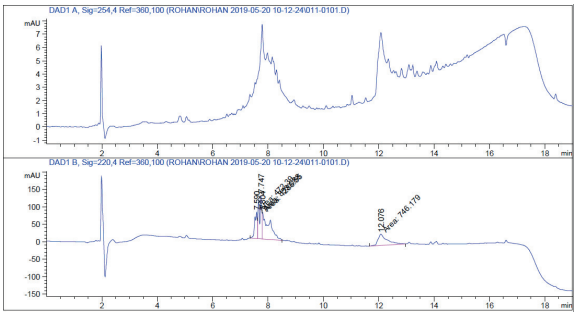
D17 unstapled: Ac-R**S$_5$**A**cBuAlaS$_5$**DISN-NH$_2$

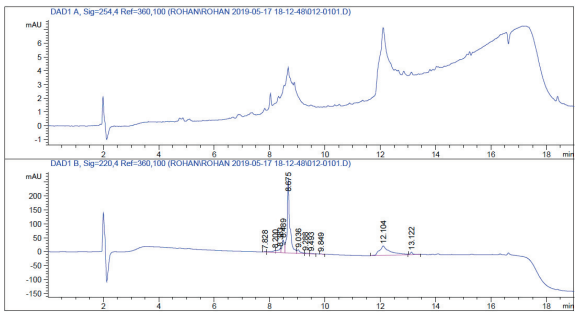


Exact Mass: 1118.64
Molecular Weight: 1119.31



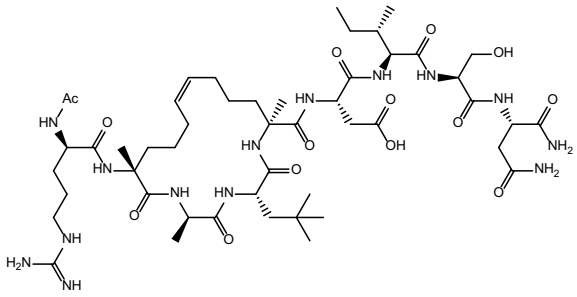




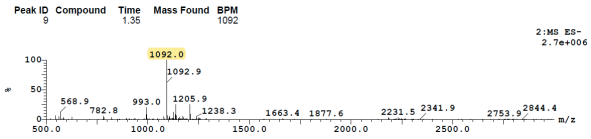


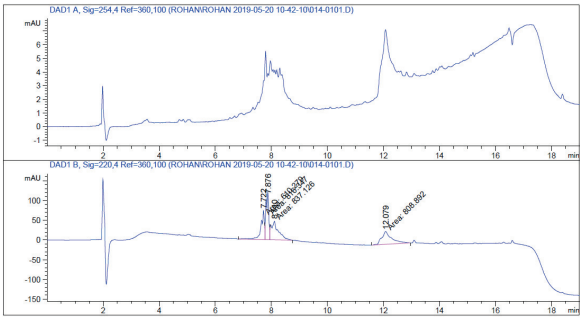
D18

D18: Ac-R**S$_5$**A**NptGlyS$_5$**DISN-NH$_2$

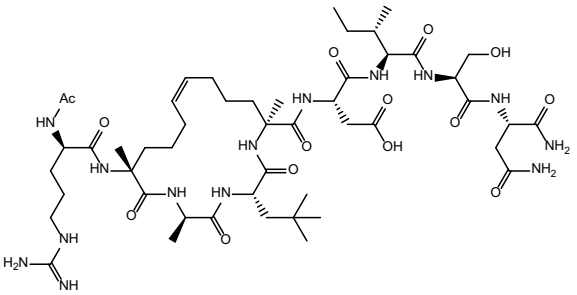


Exact Mass: 1092.63
Molecular Weight: 1093.28

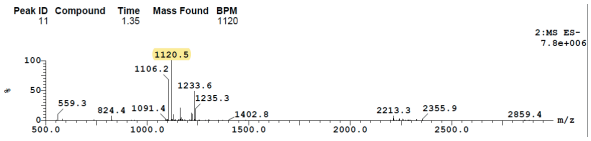


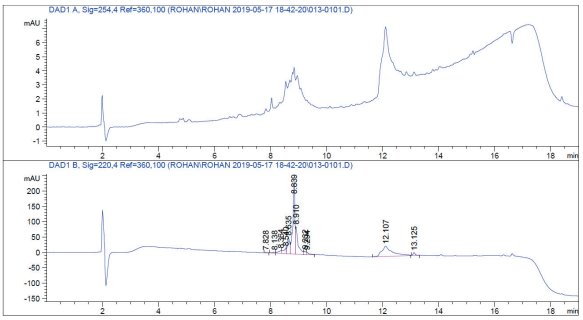


D18 unstapled: Ac-R**S$_5$**A**NptGlyS$_5$**DISN-NH$_2$



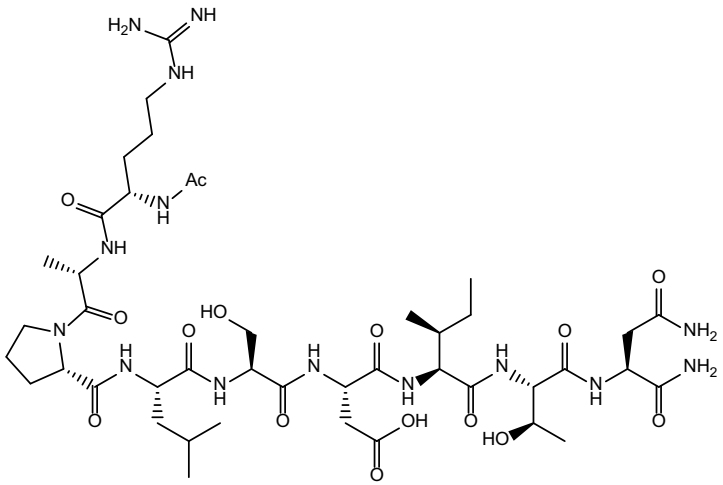
Exact Mass: 1120.66
Molecular Weight: 1121.33



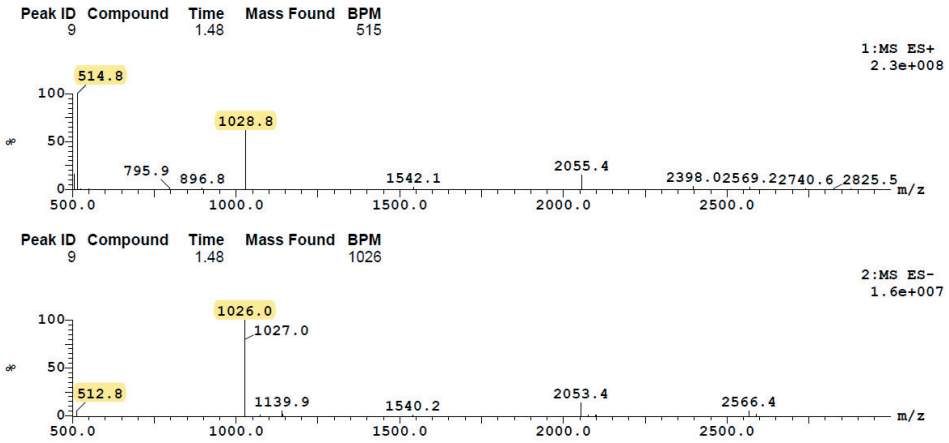


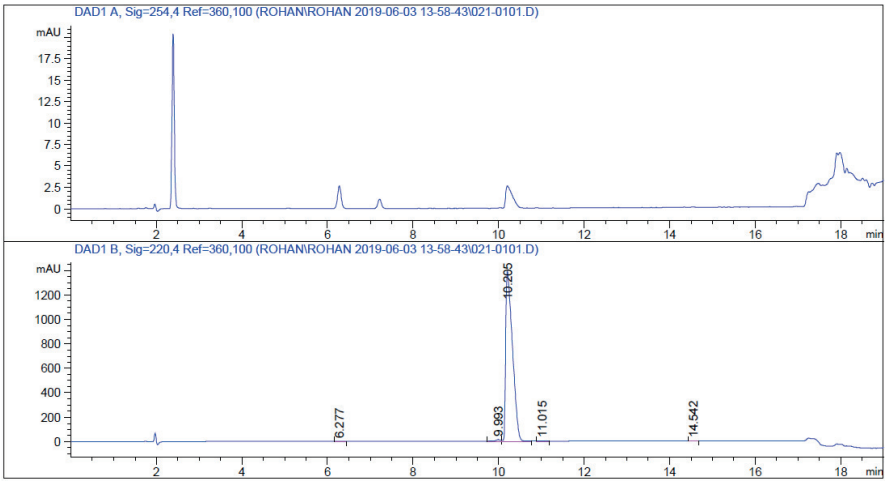
D19

D19: Ac-RAPLSDITN-NH_2



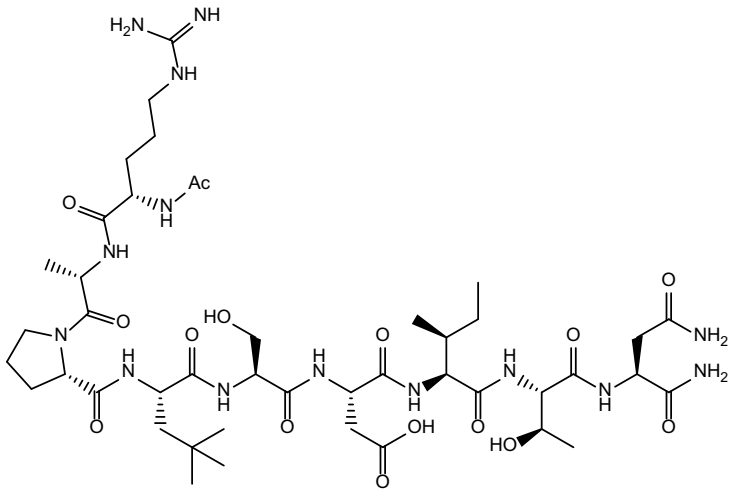
Exact Mass: 1026.55
Molecular Weight: 1027.13



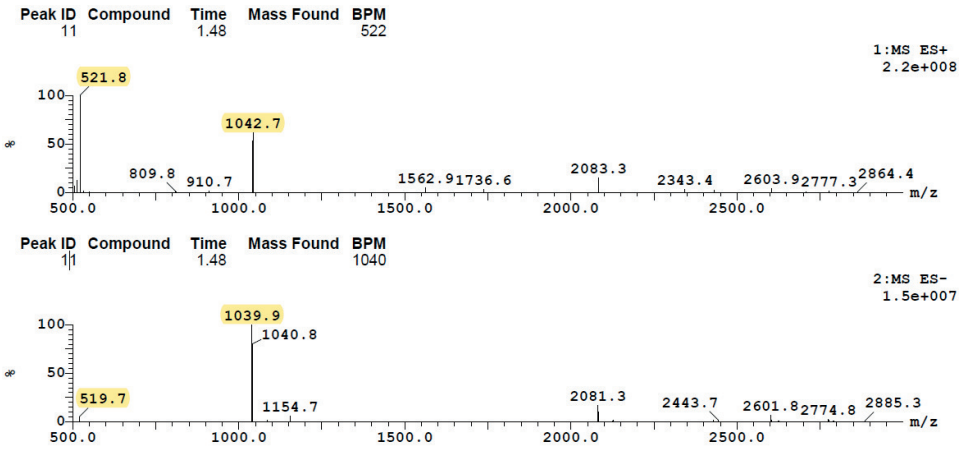


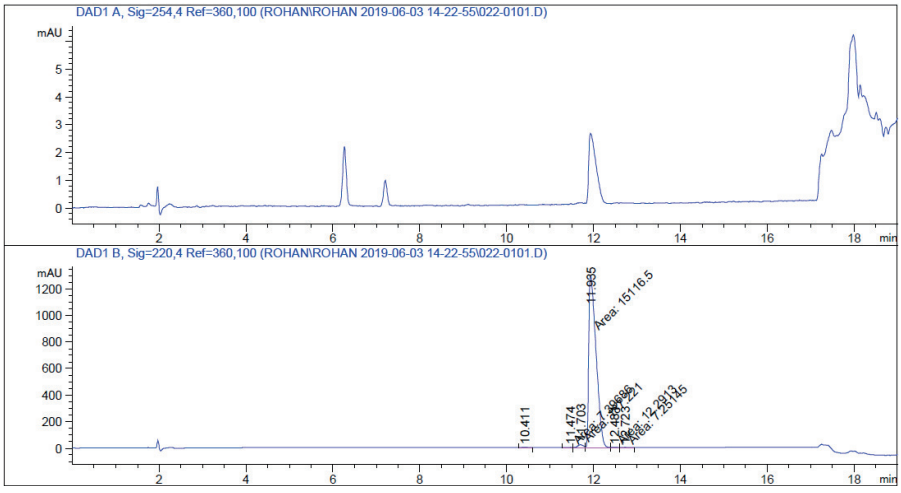
D20

D20: Ac-RAPC_3SDITN-NH_2



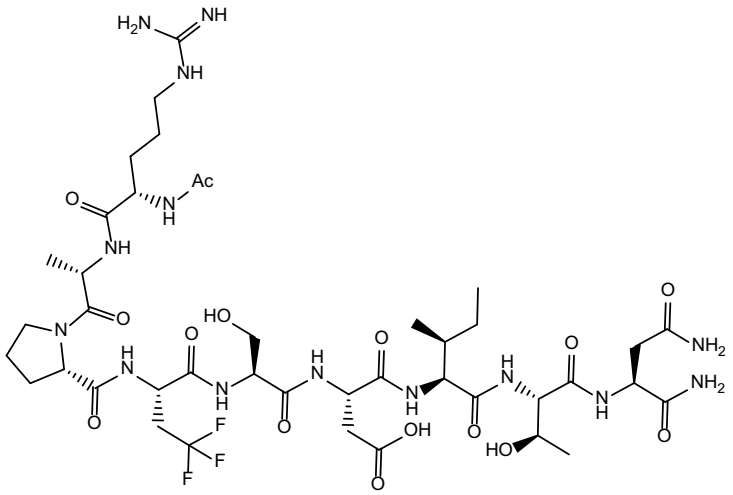
Exact Mass: 1040.56
Molecular Weight: 1041.16



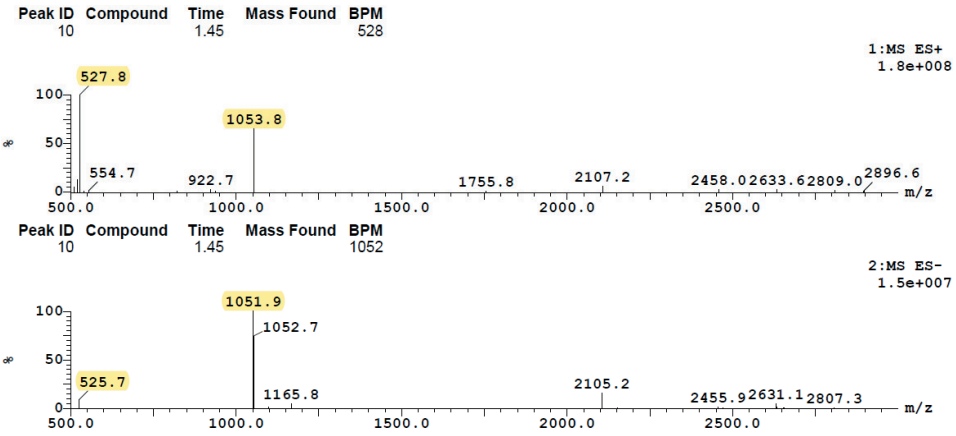


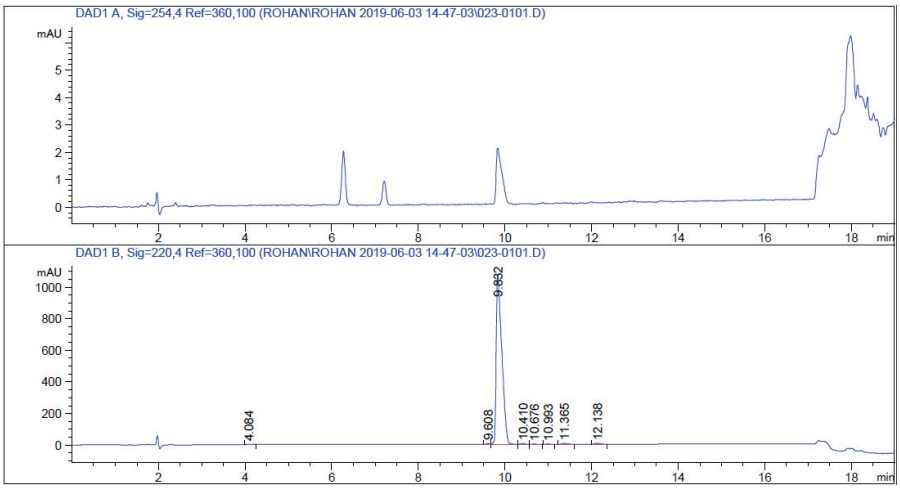
D21

D21: Ac-RAPF_3SDITN-NH_2



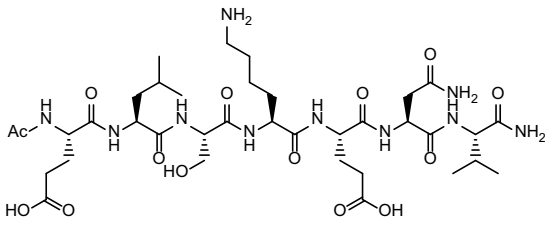
Exact Mass: 1052.49
Molecular Weight: 1053.05



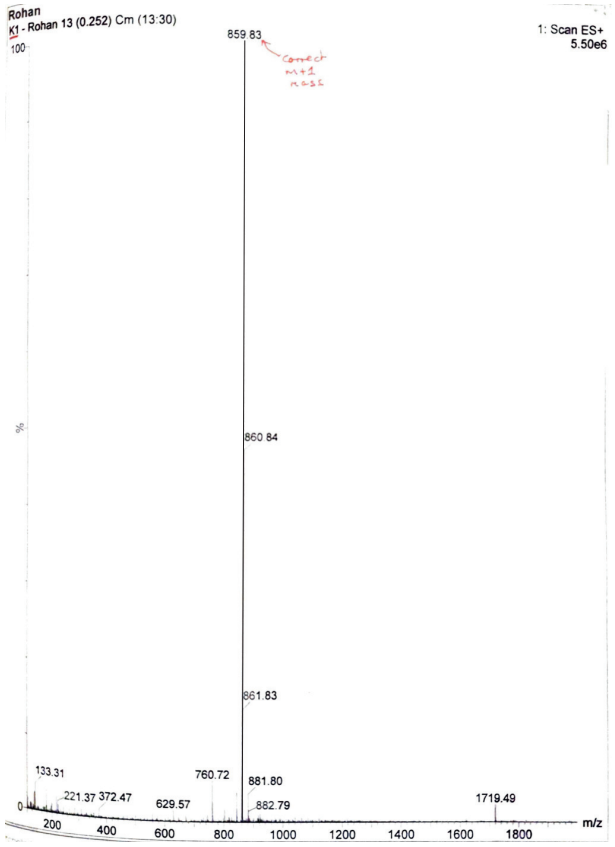


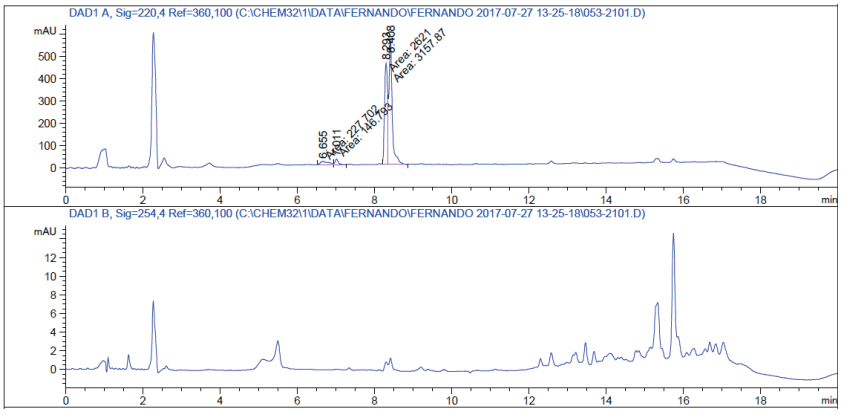
K1

K1: Ac-ELSKENV-NH_2



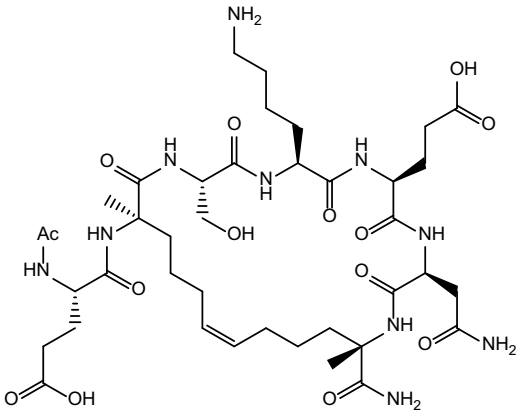
Exact Mass: 858.44
Molecular Weight: 858.94





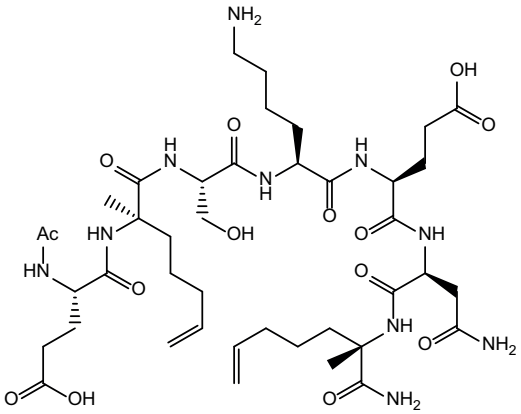
K7

K7: Ac-E$\mathbf{R_5}$SKEN$\mathbf{S_5}$-NH_2

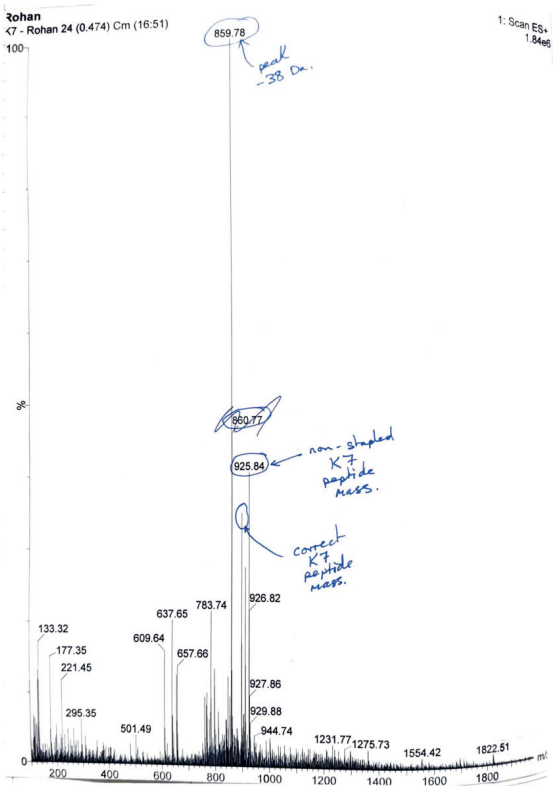


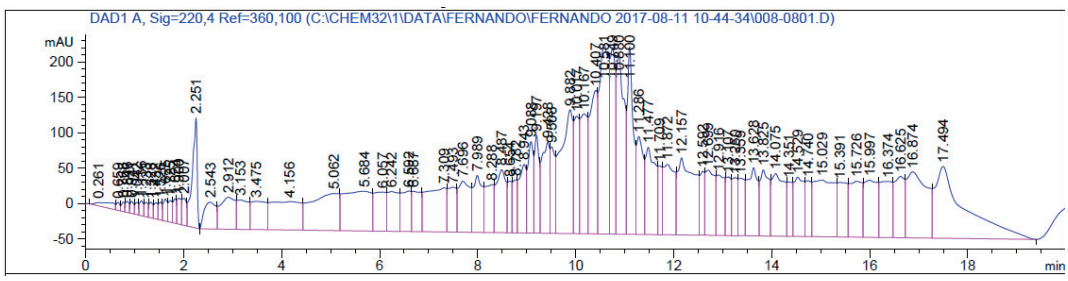
Exact Mass: 896.46
Molecular Weight: 896.98

K7 unstapled: Ac-E$\mathbf{R_5}$SKEN$\mathbf{S_5}$-NH_2



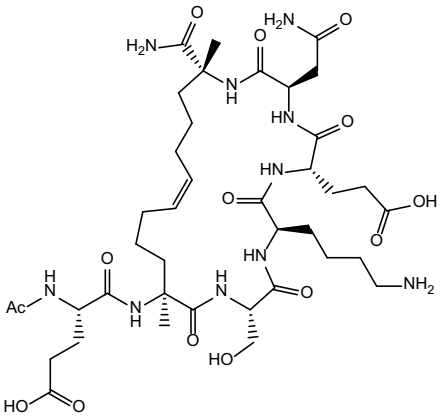
Exact Mass: 924.49
Molecular Weight: 925.04





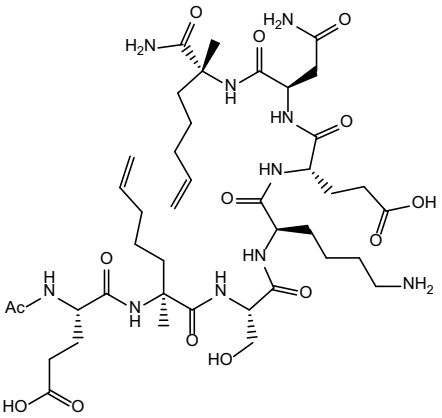
K8

K8: Ac-ES_5SKENS_5-NH_2

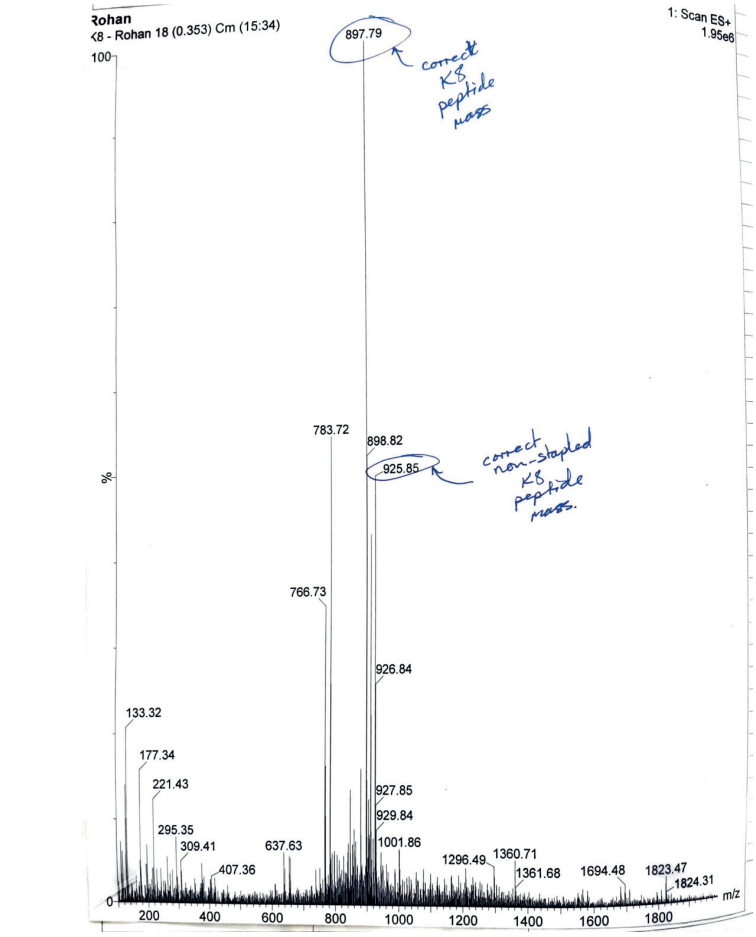


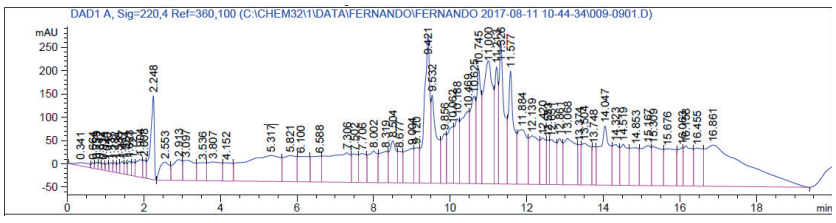
Exact Mass: 896.46
Molecular Weight: 896.98

K8 unstapled: Ac-ES_5SKENS_5-NH_2



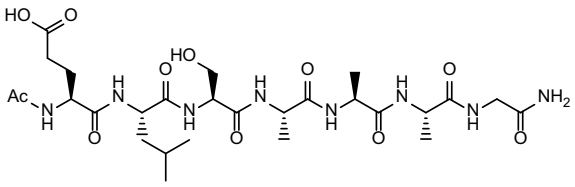
Exact Mass: 924.49
Molecular Weight: 925.04



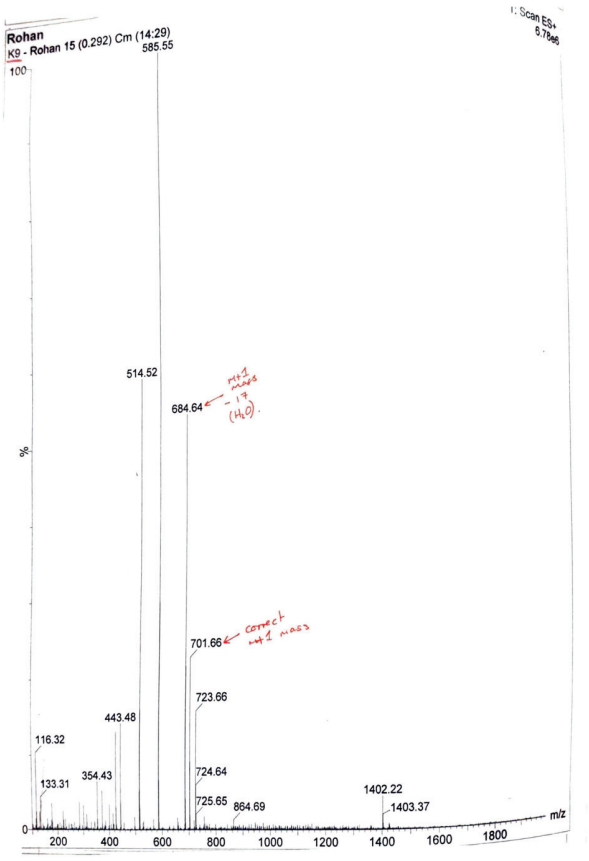


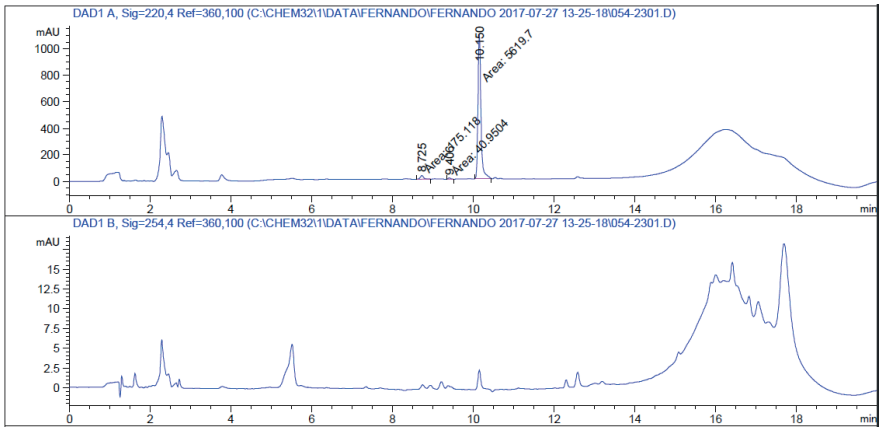
K9

K9: Ac-ELSAAAV-NH_2



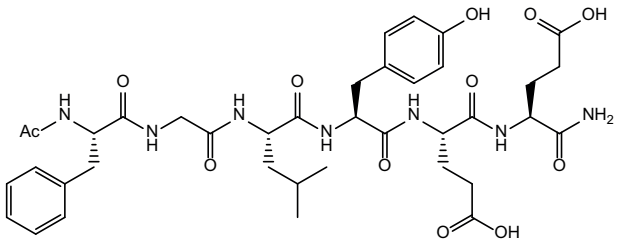
Exact Mass: 700.38
Molecular Weight: 700.78



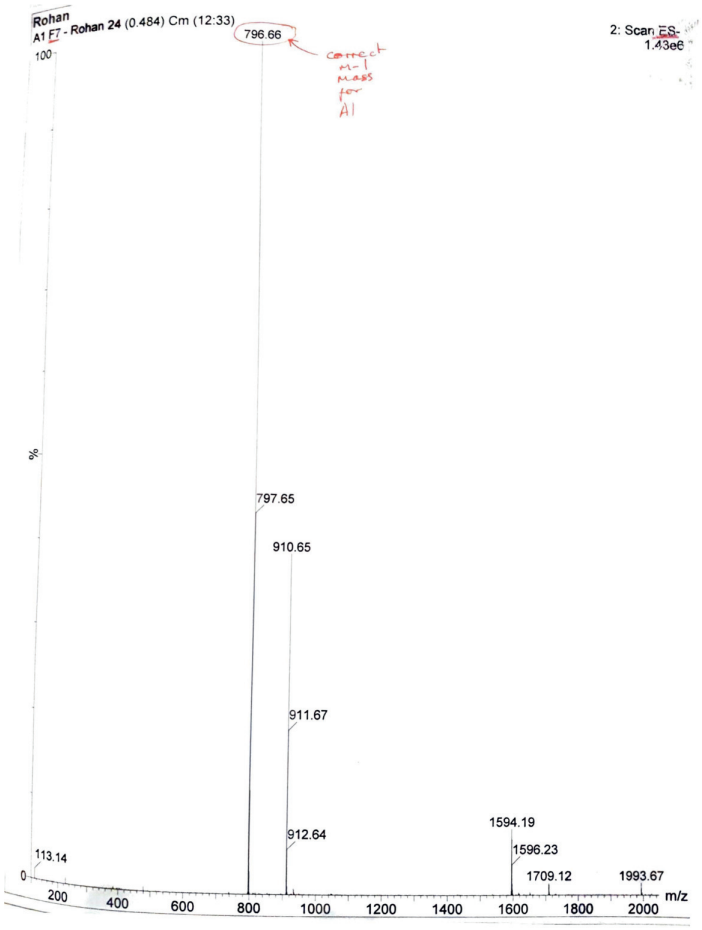


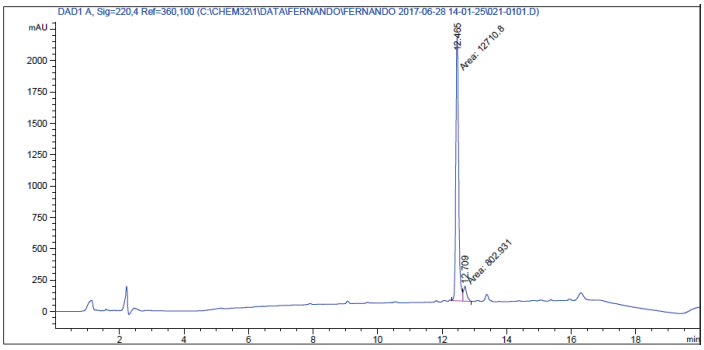
A1

A1: Ac-FGLYEE-NH_2



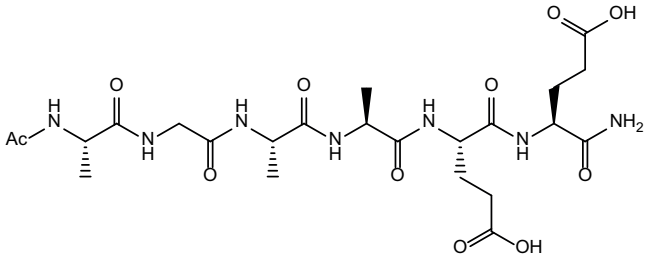
Exact Mass: 797.36
Molecular Weight: 797.85



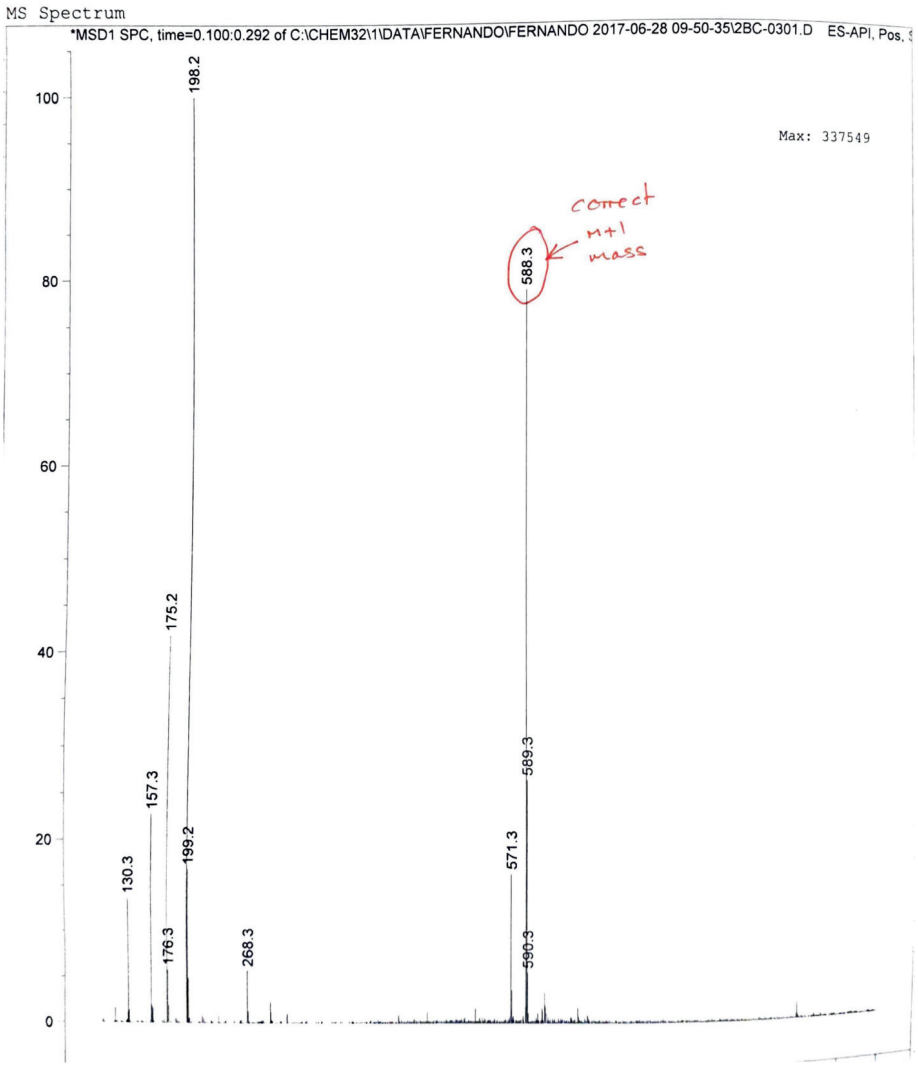


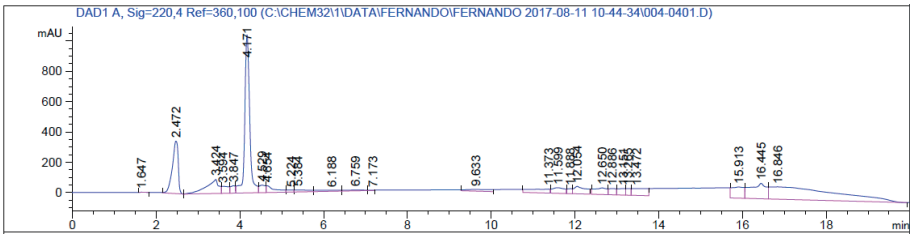
A9

A9: Ac-AGAAEE-NH_2



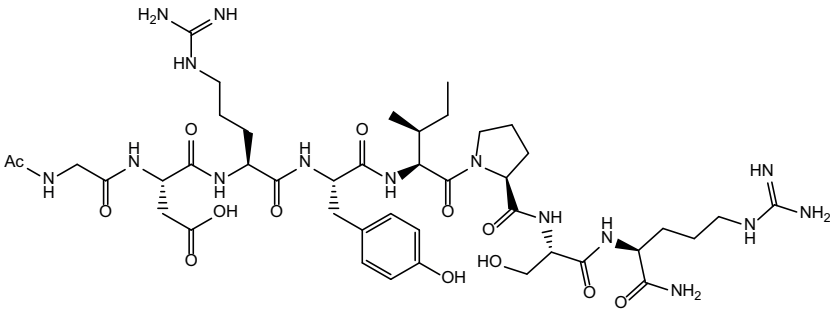
Exact Mass: 587.26
Molecular Weight: 587.58



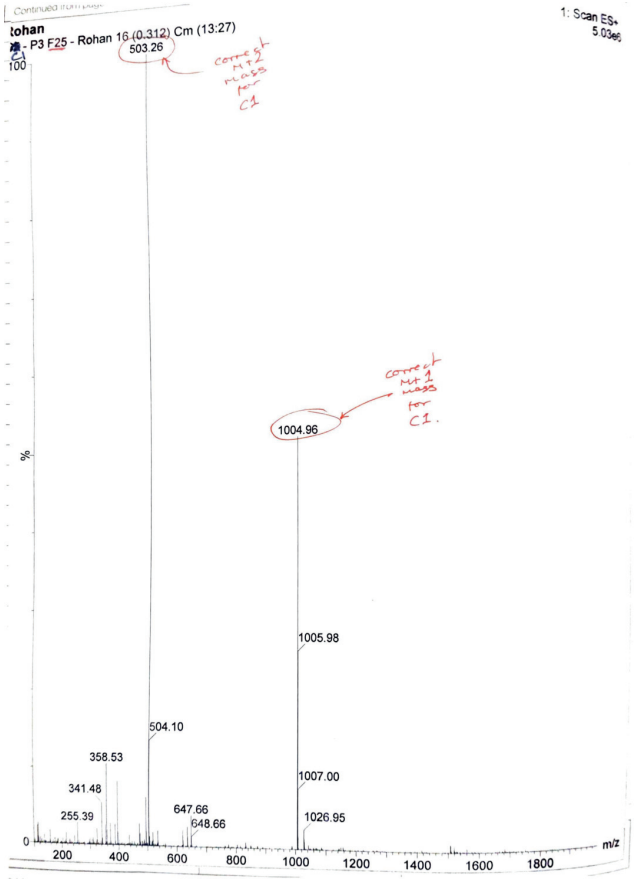


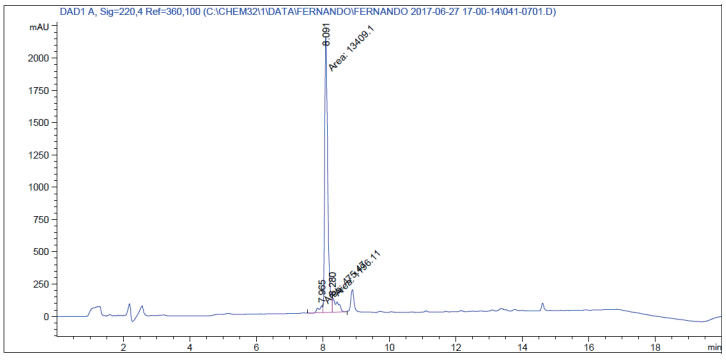
C1

C1: Ac-GDRYIPSR-NH_2



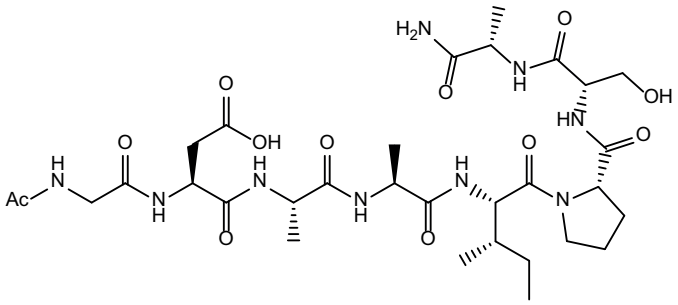
Exact Mass: 1003.52
Molecular Weight: 1004.10



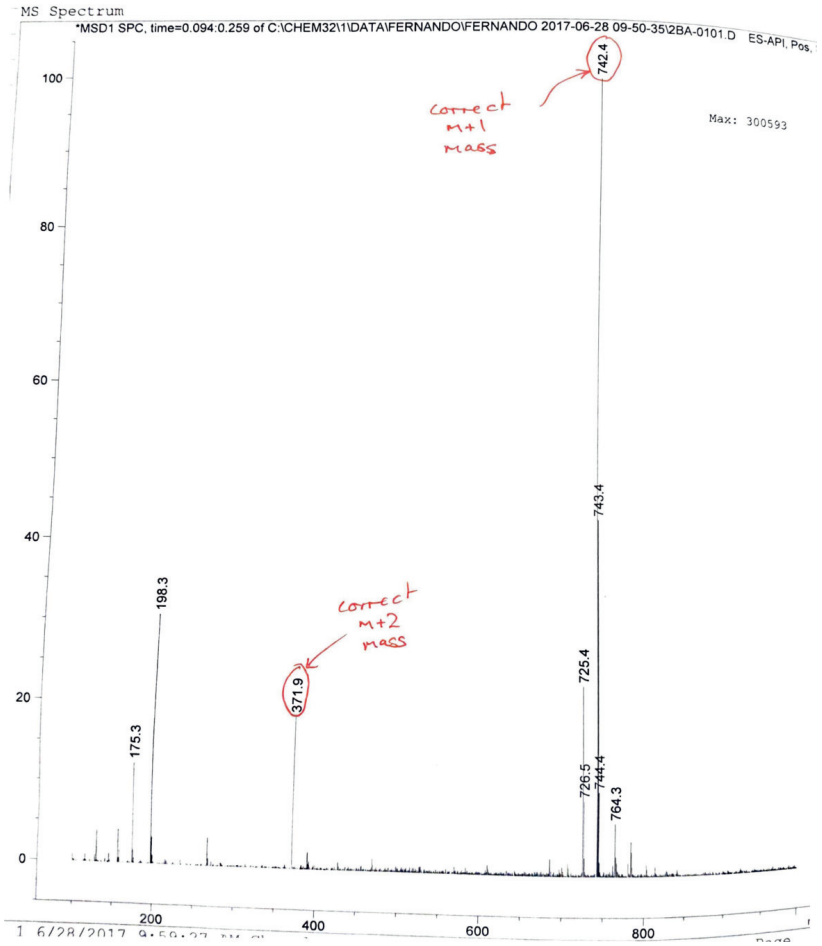


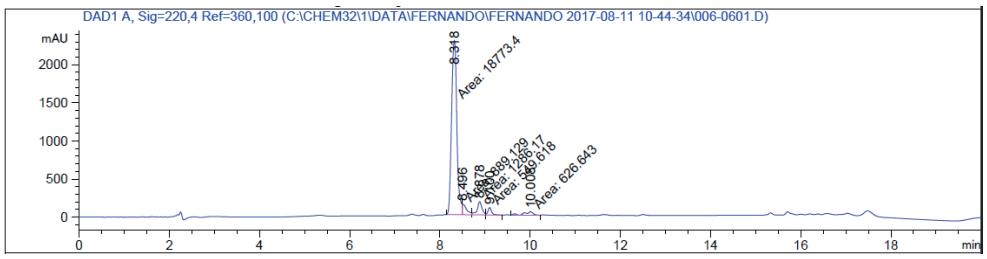
C5

C5: Ac-GDAAIPSA-NH_2



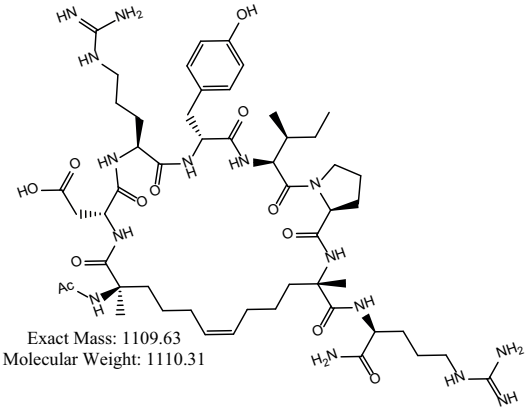
Exact Mass: 741.37
Molecular Weight: 741.79





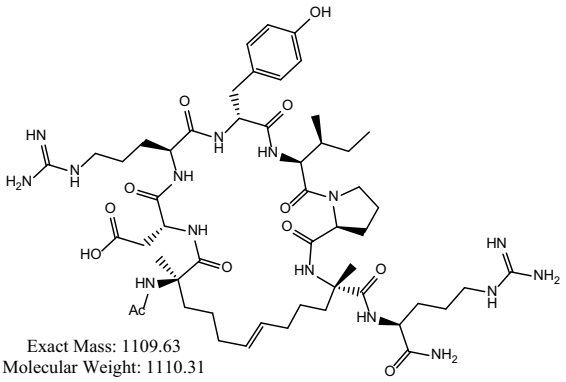
C6

C6: Ac-**S**$_5$DRYIP**S**$_5$R-NH$_2$
cis isomer

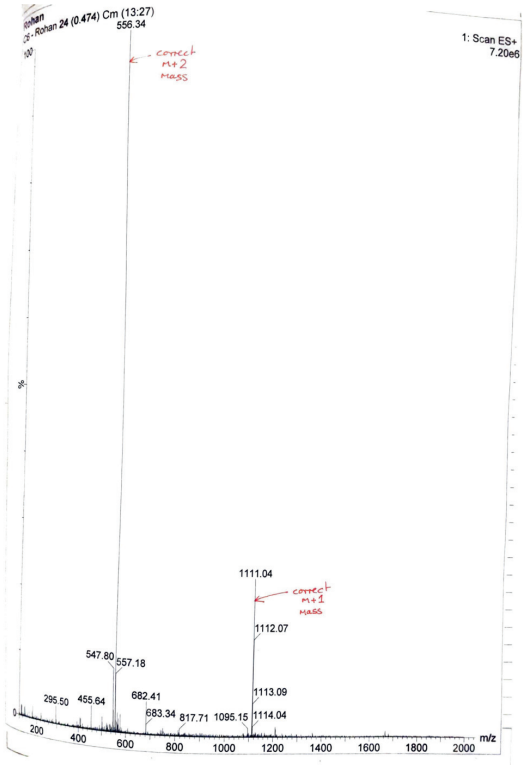


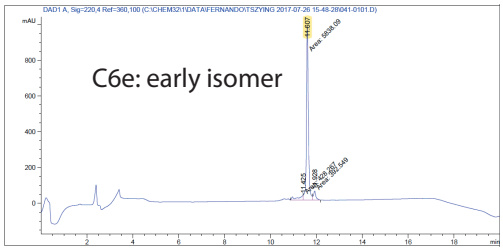
Exact Mass: 1109.63
Molecular Weight: 1110.31

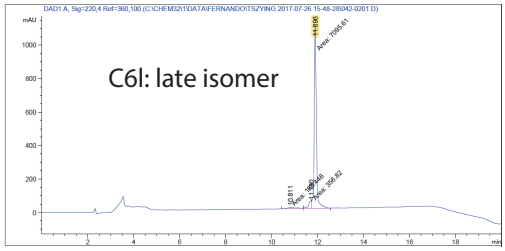
C6: Ac-**S**$_5$DRYIP**S**$_5$R-NH$_2$
trans isomer



Exact Mass: 1109.63
Molecular Weight: 1110.31







Early and late isomer are in reference to their relative retention times from analytical HPLC. The two C6 peptides isomers have not been further characterised to determine which isomer has the *cis* or *trans* configuration in the C-C double bond of the hydrocarbon staple.

Appendix C

Publications

In the attached publication by Charoenpattarapreeda *et al.* [258], Mdm2 (6-125) was purified for the collaborative work as described in Section 2.3.4. The purified protein was used in downstream high resolution ESI-MS experiments to confirm the modification of Mdm2 by P1-2 - STPOH (Figure 3). In addition, time-dependant competition fluorescence polarisation (FP) assays were performed to prove the covalent nature of their interaction with Mdm2 (Figure 4). All data were fit using equations previously described in Chapter 3, Equation 3.4. The affinity of the tracer peptide used in the assay was previously described by Lau *et al.* [96].

ChemComm

COMMUNICATION

View Article Online
View Journal | View Issue

Cite this: *Chem. Commun.*, 2019, **55**, 7914

Received 24th May 2019,
Accepted 14th June 2019

DOI: 10.1039/c9cc04022f

rsc.li/chemcomm

Targeted covalent inhibitors of MDM2 using electrophile-bearing stapled peptides†

Jiraborrirak Charoenpattarapreeda,[a] Yaw Sing Tan,[b] Jessica Iegre,[a] Stephen J. Walsh,[a] Elaine Fowler,[a] Rohan S. Eapen,[c] Yuteng Wu,[a] Hannah F. Sore,[a] Chandra S. Verma,[bde] Laura Itzhaki[c] and David R. Spring*[a]

Herein, we describe the development of a novel staple with an electrophilic warhead to enable the generation of stapled peptide covalent inhibitors of the p53–MDM2 protein–protein interaction (PPI). The peptide developed showed complete and selective covalent binding resulting in potent inhibition of p53–MDM2 PPI.

Targeted covalent inhibitors (TCIs) are a class of molecules that have been increasing in popularity and are reviving the field of covalent inhibitors.[1] TCIs employ weakly electrophilic warheads in their structure that, upon ligand binding, can react with a nucleophilic residue of the target protein. This results in a more controlled and selective binding towards the site of interest. The covalent bond only forms when the "warhead" is brought into close proximity of the appropriately located nucleophilic residue as a consequence of the inhibitor binding to the targeted pocket.[1a] Compared to non-covalent inhibitors, TCIs can offer several advantages which include the potential for improved potency,[2] longer duration of action,[3] improved selectivity, and the possibility for inhibiting 'intractable' targets.[1a,4,5]

Most covalent inhibitors have been designed to target a cysteine residue near the substrate binding site due to its low abundance in the proteome and its unique reactivity.[6] However, not all binding sites contain cysteine residues in the ideal proximity.[7] Lysine is more ubiquitous than cysteine and has been targeted in a wide variety of biological systems;[1b,7] however, examples of inhibitors targeting surface-exposed lysines are rare. The challenge for targeting the ε-NH_2 group of Lys is its high pK_{aH} (10.4) which renders 99.9% of the amino group protonated under physiological conditions.[8] Nonetheless, surface-exposed lysines can be targeted following a careful design.[1h]

MDM2, an E3 ubiquitin ligase, is a negative regulator of the tumour suppressor p53[9] and ubiquitination of p53 limits its activity.[10] Approximately 50% of human cancers possess mutated p53, whilst others feature overexpressed MDM2 resulting in the malignant cells being able to escape apoptosis.[11] Therefore, inhibition of the p53–MDM2 PPI presents a potential target for cancer therapy as documented by at least 10 compounds currently in clinical trials.[12]

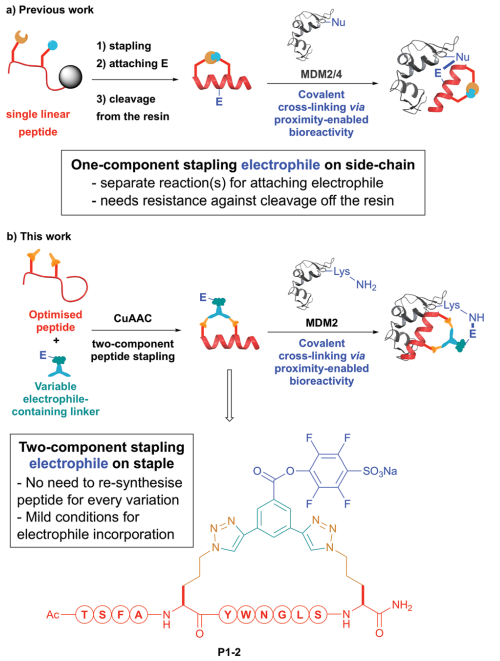
Because PPIs generally have relatively large and shallow binding pockets,[13] peptides are a suitable choice for inhibiting them due to their greater contact surface area, similar to native proteins, compared to small molecules. However, peptides on their own may suffer from poor proteolytic stability and bioavailability.[14] One of the most successful approaches to inhibit PPIs and overcome the intrinsic limitations of peptides is the use of stapled peptides – *i.e.* peptides constrained into their binding conformation by chemically cross-linking two amino acid side chains.[14a,c,d] In the case of the p53–MDM2 PPI, several stapled peptides have been developed with one example, ALRN-6924, reaching phase II clinical trials.[14b,d,15]

A recent example of a TCI is the stapled peptide mSF-SAH, which was developed to target the p53–MDM2/4 protein–protein interaction (PPI) covalently.[16] Hoppmann and Wang incorporated an unnatural amino acid with an electrophilic sulfonyl fluoride group into the peptide sequence targeting a lysine or histidine residue near the binding site on MDM2/4; one-component peptide stapling was then used to constrain the peptide (Scheme 1a).[16]

Inspired by the work of Hoppmann and Wang,[16] we propose to utilise the non-covalent binding of a stapled peptide to bring the electrophile into proximity of the targeted surface-exposed Lys residue near the MDM2 binding site. Thus, the covalent ligand-protein cross-linking is facilitated *via* proximity-enabled bioreactivity (Scheme 1b).[17] The resulting covalent linkage would prevent peptide dissociation and provide better inhibitory activity. Unlike previous reports, we envisioned to introduce the electrophilic warhead on a

[a] *Department of Chemistry, University of Cambridge, Lensfield Road, CB2 1EW, Cambridge, UK. E-mail: spring@ch.cam.ac.uk*
[b] *Bioinformatics Institute, Agency for Science, Technology and Research (A*STAR), 30 Biopolis Street, #07-01 Matrix, Singapore 138671, Singapore*
[c] *Department of Pharmacology, University of Cambridge, Tennis Court Road, Cambridge CB2 1PD, UK*
[d] *Department of Biological Sciences, National University of Singapore, 14 Science Drive 4, Singapore 117543, Singapore*
[e] *School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore 673551, Singapore*

† Electronic supplementary information (ESI) available: Supplementary figures, experimental procedures and compound characterisation data. See DOI: 10.1039/c9cc04022f



Scheme 1 (a) Previous work using one-component stapling and incorporation of an electrophilic moiety into the peptide sequence for generating a stapled peptide covalent inhibitor for MDM2/4.[16] Covalent-crosslinking occurs through proximity-enabled bioreactivity.[17] Shaded circle denotes resin. Nu = Lys or His (b) utilising 2C-PS for generation of stapled peptide covalent inhibitors. Structure of **P1-2**, the developed stapled peptide covalent inhibitor, is shown.

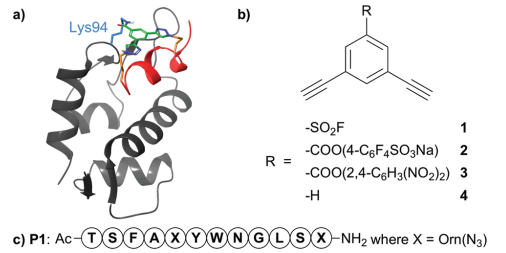


Fig. 1 (a) Molecular dynamics simulation of PMI-derived stapled peptide (red) covalently linked with MDM2 (grey) through amide bond with Lys94 close to the binding site. (b) The structures of the proposed staples: aryl sulfonyl fluoride **1**, sulfotetrafluorophenyl (STP) ester **2**, and dinitrophenyl ester **3**. The non-electrophilic staple **4** was used as a control. (c) The optimal linear diazidopeptide **P1**.

staple core compatible with the copper-catalysed azide–alkyne cycloaddition two-component peptide stapling (CuAAC 2C-PS).[14d] The use of 2C-PS improves the efficiency of the optimisation process as each variation of the staple does not require a completely new synthesis of each peptide.[18] Moreover, CuAAC 2C-PS provides mild conditions with reactions conducted at room temperature in aqueous solution and is compatible with unprotected peptides. We envisaged that by placing the electrophile on the staple and utilising a 2C-PS approach, simpler proximity tuning, staple screening, and synthesis can be achieved (Scheme 1b).

Initial analysis of the binding site of MDM2 (PDB code: 5AFG)[15d] suggested that the solvent-exposed Lys94 may be a feasible residue for covalent targeting (Fig. 1a). Molecular modelling methods were used to determine the optimal peptide sequence, staple position, and staple structure for covalent binding with Lys94.[19] We decided to adopt the phage-derived peptide PMI sequence (H-TSFAEYWNLLSP-OH) as the basis for our stapled peptide because it offers strong MDM2 binding ability and good aqueous solubility.[20] We employed the same azido-ornithine amino acid and 1,3-diethynylbenzene staple (**4**, Fig. 1b) that were used by Lau *et al.* for *i*, *i* + 7 2C-PS in our molecular models,[14b] which showed that the optimal staple positions are at residues 5 and 12, such that an electrophilic group at the 5-position of the staple could come into close contact with Lys94. Leu9 was mutated to Gly to prevent steric clash of the side chain with the staple (Fig. S1, ESI†). As a result, the peptide **P1** was identified as the optimal sequence to position the electrophilic staple in the correct location for targeting Lys94 (Fig. 1c).

We chose a sulfonyl fluoride and two activated esters for initial screening of possible electrophiles (Fig. 1b). The sulfonyl fluoride electrophile (**1**) was selected based upon previous reports of its resistance to reduction, hydrolysis, and exclusivity to heterolytic cleavage.[21] In addition, its biological application has been demonstrated through a plethora of examples including peptidic and Lys-targeted small-molecule binders.[16,22] Whereas, activated esters, namely sulfotetrafluorophenyl (STP) ester (**2**) and 2,4-dinitrophenyl ester (**3**), were chosen due to their harder electrophilic character which makes them more susceptible to an attack by an amine relative to a thiol. Moreover, activated esters are validated probes for proteomic profiling for lysine reactivity.[23] Molecular dynamics simulations showed that the ester and sulfonamide covalent complexes of the two-component stapled **P1** with Lys94 of MDM2 are stable, with the three key binding residues of the peptide (Phe3, Trp7, and Leu10) remaining bound within the p53-binding cleft at the end of all simulations (Fig. S1, ESI†).

Considering these data, we designed compounds **1**–**3** as potential electrophile-containing staples (Fig. 1b). Staples **1**, **2**, and **3** were then synthesised and tested for their compatibility with the CuAAC 2C-PS,[18] stability in aqueous media, and reactivity towards lysine. Compound **3** was insoluble in aqueous conditions; thus, it was excluded from further testing. Despite the comparable stability of staples **1** and **2** in physiological pH (49% intact for **1** *vs.* 42% for **2** after 24 h, Fig. 2a), activated ester **2** was progressed as it was found to be significantly more stable under the CuAAC 2C-PS conditions (15% intact for **1** *vs.* 83% for **2** after 24 h, Fig. 2b). In addition, **2** showed a faster reactivity with Lys whilst **1** mostly underwent hydrolysis (5% Lys-conjugated for **1** *vs.* 95% for **2** after 4 h, Fig. 2c). With these results in hand, we decided to take forward the STP ester **2** as our electrophile-containing staple. The optimised peptide **P1** was stapled with compound **2** to

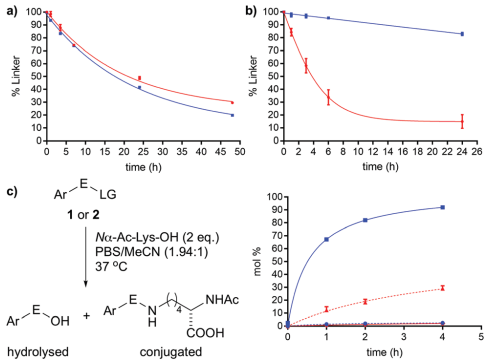


Fig. 2 Stability and reactivity tests of compounds **1** (red) and **2** (blue) (a) stability in aqueous media. (b) Stability in CuAAC condition. (c) Reactivity comparison between compounds **1** and **2** against Nα-Ac-Lys-OH. E = electrophile, LG = leaving group. Solid lines indicate conjugation while dotted lines indicate hydrolysis. All reactions were monitored by analytical HPLC and caffeine was used as the internal standard. The results are the average of two independent repeats and the errors shown as standard errors of mean.

obtain the stapled peptide **P1-2** (Scheme S1, ESI†). Pleasingly, circular dichroism measurements confirmed the staple used was able to enhance the α-helical conformation of the peptide (12% for **P1** *vs.* 22% for **P1-2**, Fig. S2, ESI†).

In order to assess the ability of **P1-2** to form a covalent complex with MDM2, **P1-2** was incubated with MDM2 and the reaction mixture analysed by mass spectrometry (Fig. 3). After the incubation, complete formation of the covalent complex was observed and no unmodified MDM2 remained (Fig. 3a and b). Pleasingly, by replacing **P1-2** with the control peptide **P1-4** (no electrophilic moiety, Scheme S1, ESI†), detection of a covalent peptide-MDM2 complex was not observed (Fig. S3, ESI†).

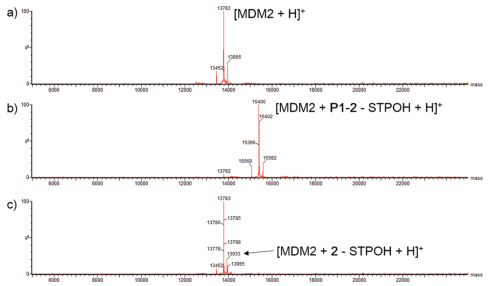
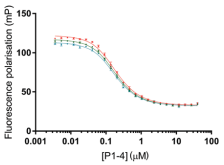


Fig. 3 ESI-MS spectra for reactions with MDM2 (a) the unmodified MDM2. $[MDM2 + H]^+$ = 13 783 Da. (b) **P1-2** (25 μM) was incubated with MDM2 (25 μM) in PBS buffer (+10% DMSO) at 37 °C for 1 h. ESI-MS spectrum of the reaction indicated the complete covalent binding of the stapled peptide. $[MDM2 + \mathbf{P1\text{-}2} - STPOH + H]^+$ = 15 400 Da (c) the same incubation as in (b) but with **P1-2** replaced with compound **2**. ESI-MS spectrum showed low reactivity of the electrophile on its own with most MDM2 unreacted. $[MDM2 + \mathbf{2} - STPOH + H]^+$ = 13 933 Da; STPOH = sulfotetrafluorophenol.

a)

Peptide	K_d (nM)†
WT PMI	16.6 ± 0.2
P1	47.8 ± 1.5
P1-4	19.3 ± 0.3

b)


c)

Time (min)	$K_{d,app}$ of **P1-2** (nM)
9	30.0 ± 9.2
41	12.9 ± 4.1
120	7.1 ± 2.1

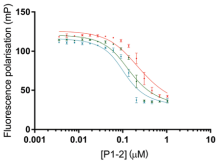
d)


Fig. 4 (a) Dissociation constants for non-covalent peptides. † The K_d values are the average of every time point and the errors are standard errors of mean. (b) Competitive FP assay of **P1-4** showing no change in affinity over time. Each curve represents one time point: 9 min (red), 41 min (green), and 120 min (blue). Each data point is arithmetic mean of triplicate and the errors shown are standard errors of mean. (c) Apparent dissociation constant of **P1-2** at each time point. (d) Competitive FP assay of **P1-2** showing the increase of the apparent dissociation constant over time which is characteristic of a covalent inhibitor.

Under the same conditions, MS results of the STP ester **2** alone (*i.e.* without peptide) with MDM2 resulted in only a small amount of the protein being modified, therefore highlighting the importance of the specific non-covalent binding of **P1-2** to MDM2 for the cross-linking event to occur (Fig. 3c).

The selectivity of the stapled peptide was investigated further by the addition of human lysozyme (Lyz) to the incubation. Lyz contains five highly solvent-accessible lysines out of which, one is particularly active.[24] Following the incubation, formation of the peptide-Lyz covalent complex was not detected indicating that **P1-2** selectively binds MDM2 (Fig. S4, ESI†).

Finally, the apparent dissociation constant ($K_{d,app}$) of **P1-2** to MDM2 was examined and compared to the non-covalent peptides using a competitive fluorescence polarisation (FP) assay (Fig. 4).[14b,15c] The linear diazidopeptide **P1** was found to have an attenuated K_d for MDM2 (K_d = 47.8 ± 1.5 nM, Fig. 4a) compared to the wild-type (WT) PMI (K_d = 16.6 ± 0.2 nM, Fig. 4a).[20] Pleasingly, the stapled peptide **P1-4** showed an affinity comparable to the WT (K_d = 19.3 ± 0.3 nM, Fig. 4a). Crucially, over the course of 120 minutes, these dissociation constants did not change significantly (Fig. 4b and Fig. S5, S6, ESI†). In contrast, as expected for a TCI, the $K_{d,app}$ of **P1-2** improved over time as the covalent bond was formed, ultimately resulting in a potent MDM2 inhibitor after 120 minutes ($K_{d,app}$ = 30.0 ± 9.2 nM at 9 min which decreased to 7.1 ± 2.1 nM at 120 min, Fig. 4c and d)‡.[25]

In summary, we have developed a novel strategy for producing stapled peptide covalent inhibitors. Expanding on our CuAAC 2C-PS technique, the staple was functionalised with a suitable electrophile for forming a covalent cross-linking with the target protein upon binding. We validated this approach using an STP ester-functionalised stapled peptide targeting the oncogenic protein MDM2: the lead peptide **P1-2** demonstrated complete covalent complex formation and nanomolar inhibition to MDM2.

Importantly, the results and concept of our study would expedite the development of stapled peptide covalent inhibitors by removing the need to synthesise the peptides anew for every sequence and requiring less demanding conditions on the electrophile. We envision that targeting proteins with low turnover rates would gain the most benefit from using this strategy.[1a,7a] A further advantage of using the CuAAC 2C-PS technique is that extra functionalisation may be achieved by appending a second functional handle to the staple and hence further enhance the capability of peptide inhibitors.

The research was supported by grants from the EPSRC, BBSRC, MRC and Royal Society. JC and JI would like to thank Trinity College, Cambridge for funding. YST and CSV would like to thank A*STAR (IAF-PP H17/01/a0/010) for support. SJW and RSE would like to thank AstraZeneca for funding.

Conflicts of interest

There are no conflicts to declare.

Notes and references

‡ Due to the instability of MDM2 in the assay, data for time-points beyond 120 minutes were not obtained.

1 (*a*) T. A. Baillie, *Angew. Chem., Int. Ed.*, 2016, **55**, 13408–13421; (*b*) M. Gehringer and S. A. Laufer, *J. Med. Chem.*, 2019, DOI: 10.1021/acs.jmedchem.8b01153; (*c*) E. Anscombe, E. Meschini, R. Mora-Vidal, S. R. Wedge, J. A. Endicott, R. J. Griffin, M. P. Martin, D. Staunton, M. Geitmann, U. H. Danielson, W. A. Stanley, L. Z. Wang, T. Reuillon, B. T. Golding, C. Cano, D. R. Newell and M. E. M. Noble, *Chem. Biol.*, 2015, **22**, 1159–1164; (*d*) S. E. Dalton, L. Dittus, D. A. Thomas, M. A. Convery, J. Nunes, J. T. Bush, J. P. Evans, T. Werner, M. Bantscheff, J. A. Murphy and S. Campos, *J. Am. Chem. Soc.*, 2018, **140**, 932–939; (*e*) T. Tamura, T. Ueda, T. Goto, T. Tsukidate, Y. Shapira, Y. Nishikawa, A. Fujisawa and I. Hamachi, *Nat. Commun.*, 2018, **9**, 1870; (*f*) M. J. Evans, A. Saghatelian, E. J. Sorensen and B. F. Cravatt, *Nat. Biotechnol.*, 2005, **23**, 1303–1307; (*g*) M. J. Evans, G. M. Morris, J. Wu, A. J. Olson, E. J. Sorensen and B. F. Cravatt, *Mol. BioSyst.*, 2007, **3**, 495; (*h*) G. Akçay, M. A. Belmonte, B. Aquila, C. Chuaqui, A. W. Hird, M. L. Lamb, P. B. Rawlins, N. Su, S. Tentarelli, N. P. Grimster and Q. Su, *Nat. Chem. Biol.*, 2016, **12**, 931–936.

2 A. J. T. Smith, X. Zhang, A. G. Leach and K. N. Houk, *J. Med. Chem.*, 2009, **52**, 225–233.

3 G. Sachs, J. M. Shin and C. W. Howden, *Aliment. Pharmacol. Ther.*, 2006, **23**, 2–8.

4 T. Barf and A. Kaptein, *J. Med. Chem.*, 2012, **55**, 6243–6262.

5 R. Lonsdale and R. A. Ward, *Chem. Soc. Rev.*, 2018, **47**, 3816–3830.

6 (*a*) S. M. Marino and V. N. Gladyshev, *J. Mol. Biol.*, 2010, **404**, 902–916; (*b*) D. A. Shannon and E. Weerapana, *Curr. Opin. Chem. Biol.*, 2015, **24**, 18–26; (*c*) S. Wu, H. Luo, H. Wang, W. Zhao, Q. Hu and Y. Yang, *Biochem. Biophys. Res. Commun.*, 2016, **478**, 1268–1273; (*d*) R. Lagoutte, R. Patouret and N. Winssinger, *Curr. Opin. Chem. Biol.*, 2017, **39**, 54–63.

7 (*a*) J. Pettinger, K. Jones and M. D. Cheeseman, *Angew. Chem., Int. Ed.*, 2017, **56**, 15200–15209; (*b*) A. Cuesta and J. Taunton, *Annu. Rev. Biochem.*, 2019, **88**, DOI: 10.1146/annurev-biochem-061516-044805.

8 G. Platzer, M. Okon and L. P. McIntosh, *J. Biomol. NMR*, 2014, **60**, 109–129.

9 A. Burgess, K. M. Chia, S. Haupt, D. Thomas, Y. Haupt and E. Lim, *Front. Oncol.*, 2016, **6**, 1–7.

10 (*a*) Y. Haupt, R. Maya, A. Kazaz and M. Oren, *Nature*, 1997, **387**, 296–299; (*b*) R. Honda, H. Tanaka and H. Yasuda, *FEBS Lett.*, 1997, **420**, 25–27.

11 (*a*) S. Shangary and S. Wang, *Clin. Cancer Res.*, 2008, **14**, 5318–5324; (*b*) F. S. Leach, T. Tokino, P. Meltzer, M. Burrell, J. D. Oliner, S. Smith, D. E. Hill, D. Sidransky, K. W. Kinzler and B. Vogelstein, *Cancer Res.*, 1993, **53**, 2231–2234; (*c*) J. D. Oliner, K. W. Kinzler, P. S. Meltzer, D. L. George and B. Vogelstein, *Nature*, 1992, **358**, 80–83.

12 (*a*) V. Tisato, R. Voltan, A. Gonelli, P. Secchiero and G. Zauli, *J. Hematol. Oncol.*, 2017, **10**, 133; (*b*) D. Nguyen, W. Liao, S. X. Zeng and H. Lu, *Pharmacol. Ther.*, 2017, **178**, 92–108; (*c*) A. Aguilar, J. Lu, L. Liu, D. Du, D. Bernard, D. McEachern, S. Przybranowski, X. Li, R. Luo, B. Wen, D. Sun, H. Wang, J. Wen, G. Wang, Y. Zhai, M. Guo, D. Yang and S. Wang, *J. Med. Chem.*, 2017, **60**, 2819–2839.

13 (*a*) D. E. Scott, A. R. Bayly, C. Abell and J. Skidmore, *Nat. Rev. Drug Discovery*, 2016, **15**, 533–550; (*b*) J. A. Wells and C. L. McClendon, *Nature*, 2007, **450**, 1001–1009.

14 (*a*) L. D. Walensky and G. H. Bird, *J. Med. Chem.*, 2014, **57**, 6275–6288; (*b*) Y. H. Lau, P. de Andrade, S.-T. Quah, M. Rossmann, L. Laraia, N. Sköld, T. J. Sum, P. J. E. Rowling, T. L. Joseph, C. Verma, M. Hyvönen, L. S. Itzhaki, A. R. Venkitaraman, C. J. Brown, D. P. Lane and D. R. Spring, *Chem. Sci.*, 2014, **5**, 1804–1809; (*c*) Y. H. Lau, P. de Andrade, Y. Wu and D. R. Spring, *Chem. Soc. Rev.*, 2015, **44**, 91–102; (*d*) J. Iegre, J. S. Gaynord, N. S. Robertson, H. F. Sore, M. Hyvönen and D. R. Spring, *Adv. Ther.*, 2018, **1**, 1800052.

15 (*a*) F. Bernal, A. F. Tyler, S. J. Korsmeyer, L. D. Walensky and G. L. Verdine, *J. Am. Chem. Soc.*, 2007, **129**, 2456–2457; (*b*) Y. S. Chang, B. Graves, V. Guerlavais, C. Tovar, K. Packman, K.-H. To, K. A. Olson, K. Kesavan, P. Gangurde, A. Mukherjee, T. Baker, K. Darlak, C. Elkin, Z. Filipovic, F. Z. Qureshi, H. Cai, P. Berry, E. Feyfant, X. E. Shi, J. Horstick, D. A. Annis, A. M. Manning, N. Fotouhi, H. Nash, L. T. Vassilev and T. K. Sawyer, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, E3445–E3454; (*c*) C. J. Brown, S. T. Quah, J. Jong, A. M. Goh, P. C. Chiam, K. H. Khoo, M. L. Choong, M. A. Lee, L. Yurlova, K. Zolghadr, T. L. Joseph, C. S. Verma and D. P. Lane, *ACS Chem. Biol.*, 2013, **8**, 506–512; (*d*) Y. H. Lau, Y. Wu, M. Rossmann, B. X. Tan, P. de Andrade, Y. S. Tan, C. Verma, G. J. McKenzie, A. R. Venkitaraman, M. Hyvönen and D. R. Spring, *Angew. Chem., Int. Ed.*, 2015, **54**, 15410–15413; (*e*) K. Hu, F. Yin, M. Yu, C. Sun, J. Li, Y. Liang, W. Li, M. Xie, Y. Lao, W. Liang and Z.-G. Li, *Theranostics*, 2017, **7**, 4566–4576; (*f*) X. Li, W. D. Tolbert, H.-G. Hu, N. Gohain, Y. Zou, F. Niu, W.-X. He, W. Yuan, J.-C. Su, M. Pazgier and W. Lu, *Chem. Sci.*, 2019, **10**, 1522–1530.

16 C. Hoppmann and L. Wang, *Chem. Commun.*, 2016, **52**, 5140–5143.

17 Z. Xiang, H. Ren, Y. S. Hu, I. Coin, J. Wei, H. Cang and L. Wang, *Nat. Methods*, 2013, **10**, 885–888.

18 Y. H. Lau, Y. Wu, P. De Andrade, W. R. J. D. Galloway and D. R. Spring, *Nat. Protoc.*, 2015, **10**, 585–594.

19 Y. S. Tan, D. P. Lane and C. S. Verma, *Drug Discovery Today*, 2016, **21**, 1642–1653.

20 M. Pazgier, M. Liu, G. Zou, W. Yuan, C. Li, C. Li, J. Li, J. Monbo, D. Zella, S. G. Tarasov and W. Lu, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 4665–4670.

21 J. Dong, L. Krasnova, M. G. Finn and K. Barry Sharpless, *Angew. Chem., Int. Ed.*, 2014, **53**, 9430–9448.

22 (*a*) N. N. Gushwa, S. Kang, J. Chen and J. Taunton, *J. Am. Chem. Soc.*, 2012, **134**, 20214–20217; (*b*) A. J. Brouwer, A. Jonker, P. Werkhoven, E. Kuo, N. Li, N. Gallastegui, J. Kemmink, B. I. Florea, M. Groll, H. S. Overkleeft and R. M. J. Liskamp, *J. Med. Chem.*, 2012, **55**, 10995–11003; (*c*) N. P. Grimster, S. Connelly, A. Baranczak, J. Dong, L. B. Krasnova, K. B. Sharpless, E. T. Powers, I. A. Wilson and J. W. Kelly, *J. Am. Chem. Soc.*, 2013, **135**, 5656–5668; (*d*) Q. Zhao, X. Ouyang, X. Wan, K. S. Gajiwala, J. C. Kath, L. H. Jones, A. L. Burlingame and J. Taunton, *J. Am. Chem. Soc.*, 2017, **139**, 680–685.

23 S. M. Hacker, K. M. Backus, M. R. Lazear, S. Forli, B. E. Correia and B. F. Cravatt, *Nat. Chem.*, 2017, **9**, 1181–1190.

24 M. Ahn, E. De Genst, G. S. Kaminski Schierle, M. Erdelyi, C. F. Kaminski, C. M. Dobson and J. R. Kumita, *PLoS One*, 2012, **7**, e50192.

25 J. Singh, R. C. Petter, T. A. Baillie and A. Whitty, *Nat. Rev. Drug Discovery*, 2011, **10**, 307–317.

Bibliography

[1] Androulla N. Miliotou and Lefkothea C. Papadopoulou. CAR T-cell Therapy: A New Era in Cancer Immunotherapy. *Curr. Pharm. Biotechnol.*, 19(1):5–18, 2018.

[2] Adrian Pickar-Oliver and Charles A. Gersbach. The next generation of CRISPR–Cas technologies and applications. *Nat. Rev. Mol. Cell Biol.*, 20(8):490–507, 2019.

[3] J.K. Borchardt. The beginnings of drug therapy: Ancient mesopotamian medicine. *Drug News Perspect.*, 2002.

[4] Roswell Quinn. Rethinking antibiotic research and development: World War II and the penicillin collaborative. *Am. J. Public Health*, 103(3):426–434, 2013.

[5] A G Fleming. Responsibilities and Opportunities of the Private Practitioner in Preventive Medicine. *Can. Med. Assoc. J.*, 20(1):11–3, 1929.

[6] Albert Schatz, Elizabeth Bugle, and Selman A. Waksman. Streptomycin, a Substance Exhibiting Antibiotic Activity Against Gram-Positive and Gram-Negative Bacteria. *Proc. Soc. Exp. Biol. Med.*, 55(1):66–69, 1944.

[7] Neil Vargesson. Thalidomide-induced teratogenesis: History and mechanisms. *Birth Defects Res. Part C - Embryo Today Rev.*, 105(2):140–156, 2015.

[8] Milestones in U.S. Food and Drug Law History | FDA. https://www.fda.gov/about-fda/fdas-evolving-regulatory-powers/milestones-us-food-and-drug-law-history.

[9] Matthew Cobb. 60 years ago, Francis Crick changed the logic of biology. *PLoS Biol.*, 15(9):e2003243, 2017.

[10] Takumi Ito, Hideki Ando, Takayuki Suzuki, Toshihiko Ogura, Kentaro Hotta, Yoshimasa Imamura, Yuki Yamaguchi, and Hiroshi Handa. Identification of a primary target of thalidomide teratogenicity. *Science (80-.).*, 327(5971):1345–1350, 2010.

[11] AstraZeneca Annual report 2019. https://www.astrazeneca.com/investor-relations/annual-reports/annual-report-2019.html.

[12] AstraZeneca Research and Development Expenses 2006-2020 | AZN | MacroTrends. https://www.macrotrends.net/stocks/charts/AZN/astrazeneca/research-development-expenses.

[13] Jack W. Scannell, Alex Blanckley, Helen Boldon, and Brian Warrington. Diagnosing the decline in pharmaceutical R&D efficiency. *Nat. Rev. Drug Discov.*, 11(3):191–

200, 2012.

[14] Jack W. Scannell and Jim Bosley. When quality beats quantity: Decision theory, drug discovery, and the reproducibility crisis. *PLoS One*, 11(2):e0147215, 2016.

[15] Michael B. Bracken. Why animal studies are often poor predictors of human reactions to exposure. *J. R. Soc. Med.*, 102(3):120–122, 2009.

[16] Monica J. Justice and Paraminder Dhillon. Using the mouse to model human disease: Increasing validity and reproducibility. *DMM Dis. Model. Mech.*, 9(2): 101–103, 2016.

[17] Philip Cohen. Protein kinases - The major drug targets of the twenty-first century? *Nat. Rev. Drug Discov.*, 1(4):309–315, 2002.

[18] Khushwant S. Bhullar, Naiara Orrego Lagarón, Eileen M. McGowan, Indu Parmar, Amitabh Jha, Basil P. Hubbard, and H. P.Vasantha Rupasinghe. Kinase-targeted cancer therapies: Progress, challenges and future directions. *Mol. Cancer*, 17(48), 2018.

[19] Jéôrme Hert, John J. Irwin, Christian Laggner, Michael J. Keiser, and Brian K. Shoichet. Quantifying biogenic bias in screening libraries. *Nat. Chem. Biol.*, 5(7): 479–483, 2009.

[20] J. J. Ward, J. S. Sodhi, L. J. McGuffin, B. F. Buxton, and D. T. Jones. Prediction and Functional Analysis of Native Disorder in Proteins from the Three Kingdoms of Life. *J. Mol. Biol.*, 337(3):635–645, 2004.

[21] Robin Van Der Lee, Marija Buljan, Benjamin Lang, Robert J. Weatheritt, Gary W. Daughdrill, A. Keith Dunker, Monika Fuxreiter, Julian Gough, Joerg Gsponer, David T. Jones, Philip M. Kim, Richard W. Kriwacki, Christopher J. Oldfield, Rohit V. Pappu, Peter Tompa, Vladimir N. Uversky, Peter E. Wright, and M. Madan Babu. Classification of intrinsically disordered regions and proteins. *Chem. Rev.*, 114(13):6589–6631, 2014.

[22] Mark Wells, Henning Tidow, Trevor J. Rutherford, Phineus Markwick, Malene Ringkjobing Jensen, Efstratios Mylonas, Dmitri I. Svergun, Martin Blackledge, and Alan R. Fersht. Structure of tumor suppressor p53 and its intrinsically disordered N-terminal transactivation domain. *Proc. Natl. Acad. Sci. U. S. A.*, 105(15): 5762–5767, 2008.

[23] Vladimir N. Uversky. p53 proteoforms and intrinsic disorder: An illustration of the protein structure–function continuum concept. *Int. J. Mol. Sci.*, 17(1874), 2016.

[24] Carl W. Anderson and Ettore Appella. Signaling to the p53 tumor suppressor through pathways activated by genotoxic and non-genotoxic stresses. In *Handb. Cell Signaling, 2/e*, volume 3, pages 2185–2204. Elsevier Inc., 2010.

[25] Nicolas Basse, Joel L. Kaar, Giovanni Settanni, Andreas C. Joerger, Trevor J.

Rutherford, and Alan R. Fersht. Toward the Rational Design of p53-Stabilizing Drugs: Probing the Surface of the Oncogenic Y220C Mutant. *Chem. Biol.*, 17(1): 46–56, 2010.

[26] Andreas C. Joerger and Alan R. Fersht. The tumor suppressor p53: from structures to drug discovery. *Cold Spring Harb. Perspect. Biol.*, 2(6):a000919, 2010.

[27] Mary E. Matyskiela and David O. Morgan. Analysis of Activator-Binding Sites on the APC/C Supports a Cooperative Substrate-Binding Mechanism. *Mol. Cell*, 34 (1):68–80, 2009.

[28] Suyang Zhang, Leifu Chang, Claudio Alfieri, Ziguo Zhang, Jing Yang, Sarah Maslen, Mark Skehel, and David Barford. Molecular mechanism of APC/C activation by mitotic phosphorylation. *Nature*, 533:260–264, 2016.

[29] Duncan E. Scott, Andrew R. Bayly, Chris Abell, and John Skidmore. Small molecules, big targets: Drug discovery faces the protein-protein interaction challenge. *Nat. Rev. Drug Discov.*, 15(8):533–550, 2016.

[30] Eugene Valkov, Tim Sharpe, May Marsh, Sandra Greive, and Marko Hyvönen. Targeting Protein–Protein Interactions and Fragment-Based Drug Discovery. In *Top. Curr. Chem.*, pages 145–179. Springer, Berlin, Heidelberg, 2011.

[31] Michelle R. Arkin, Yinyan Tang, and James A. Wells. Small-molecule inhibitors of protein-protein interactions: Progressing toward the reality. *Chem. Biol.*, 21(9): 1102–1114, 2014.

[32] Sivaraman Dandapani, Gerard Rosse, Noel Southall, Joseph M. Salvino, and Craig J. Thomas. Selecting, Acquiring, and Using Small Molecule Libraries for High-Throughput Screening. *Curr. Protoc. Chem. Biol.*, 4(3):177–191, 2012.

[33] Anne Mai Wassermann, Luiz Miguel Camargo, and Douglas S. Auld. Composition and applications of focus libraries to phenotypic assays. *Front. Pharmacol.*, 5 JUN: 164, 2014.

[34] Lyubomir T. Vassilev, Binh T. Vu, Bradford Graves, Daisy Carvajal, Frank Podlaski, Zoran Filipovic, Norman Kong, Ursula Kammlott, Christine Lukacs, Christian Klein, Nader Fotouhi, and Emily A. Liu. In Vivo Activation of the p53 Pathway by Small-Molecule Antagonists of MDM2. *Science (80-.).*, 303(5659):844–848, 2004.

[35] Panagis Filippakopoulos, Jun Qi, Sarah Picaud, Yao Shen, William B. Smith, Oleg Fedorov, Elizabeth M. Morse, Tracey Keates, Tyler T. Hickman, Ildiko Felletar, Martin Philpott, Shonagh Munro, Michael R. McKeown, Yuchuan Wang, Amanda L. Christie, Nathan West, Michael J. Cameron, Brian Schwartz, Tom D. Heightman, Nicholas La Thangue, Christopher A. French, Olaf Wiest, Andrew L. Kung, Stefan Knapp, and James E. Bradner. Selective inhibition of BET bromodomains. *Nature*, 468(7327):1067–1073, 2010.

[36] Amanda J. Price, Steven Howard, and Benjamin D. Cons. Fragment-based drug discovery and its application to challenging drug targets. *Essays Biochem.*, 61(5): 475–484, 2017.

[37] Kenneth Hallenbeck, David Turner, Adam Renslo, and Michelle Arkin. Targeting Non-Catalytic Cysteine Residues Through Structure-Guided Drug Discovery. *Curr. Top. Med. Chem.*, 17(1):4–15, 2016.

[38] Jolene L. Lau and Michael K. Dunn. Therapeutic peptides: Historical perspectives, current development trends, and future directions. *Bioorganic Med. Chem.*, 26(10): 2700–2707, 2018.

[39] David J. Newman and Gordon M. Cragg. Bioactive macrocycles from nature. *RSC Drug Discov. Ser.*, 2015(40):1–36, 2015.

[40] Eric Marsault and Mark L. Peterson. Macrocycles are great cycles: Applications, opportunities, and challenges of synthetic macrocycles in drug discovery. *J. Med. Chem.*, 54(7):1961–2004, 2011.

[41] Patrick G. Dougherty, Ziqing Qian, and Dehua Pei. Macrocycles as protein-protein interaction inhibitors. *Biochem. J.*, 474(7):1109–1125, 2017.

[42] Nicholas K. Terrett. Methods for the synthesis of macrocycle libraries for drug discovery. *Drug Discov. Today Technol.*, 7(2):e94–e104, 2010.

[43] Dmitry L. Usanov, Alix I. Chan, Juan Pablo Maianti, and David R. Liu. Second-generation DNA-templated macrocycle libraries for the discovery of bioactive small molecules. *Nat. Chem.*, 10(7):704–714, 2018.

[44] Jing Li, Sang Gyun Kim, and John Blenis. Rapamycin: One drug, many effects. *Cell Metab.*, 19(3):373–379, 2014.

[45] F. G. Banting, C. H. Best, J. B. Collip, W. R. Campbell, and A. A. Fletcher. Pancreatic extracts in the treatment of diabetes mellitus. *Indian J. Med. Res.*, 125 (3):141–146, 1922.

[46] J. R. Elkinton, A. D. Hunt, L. Godfrey, W. W. Mccrory, A. G. Rogerson, and J. Stokes. Effects of pituitary adrenocorticotropic hormone (acth) therapy. *J. Am. Med. Assoc.*, 141(18):1273–1279, 1949.

[47] D. H. Copp and B. Cheney. Calcitonin - A hormone from the parathyroid which lowers the calcium-level of the blood. *Nature*, 193(4813):381–382, 1962.

[48] M. P. Embrey and J. F. Anselmo. The effects of intravenous oxytocin on uterine contractility. *BJOG An Int. J. Obstet. Gynaecol.*, 69(6):918–923, 1962.

[49] M. A. Warso, J. M. Richards, D. Mehta, K. Christov, C. Schaeffer, L. Rae Bressler, T. Yamada, D. Majumdar, S. A. Kennedy, C. W. Beattie, and T. K. Das Gupta. A first-in-class, first-in-human, phase i trial of p28, a non-HDM2-mediated peptide inhibitor of p53 ubiquitination in patients with advanced solid tumours. *Br. J.*

Cancer, 108(5):1061–1070, 2013.

[50] Rishi R. Lulla, Stewart Goldman, Tohru Yamada, Craig W. Beattie, Linda Bressler, Michael Pacini, Ian F. Pollack, Paul Graham Fisher, Roger J. Packer, Ira J. Dunkel, Girish Dhall, Shengjie Wu, Arzu Onar, James M. Boyett, and Maryam Fouladi. Phase i trial of p28 (NSC745104), a non-HDM2-mediated peptide inhibitor of p53 ubiquitination in pediatric patients with recurrent or progressive central nervous system tumors: A Pediatric Brain Tumor Consortium Study. *Neuro. Oncol.*, 18(9): 1319–1325, 2016.

[51] Funda Meric-Bernstam, Mansoor N. Saleh, Jeffrey R. Infante, Sanjay Goel, Gerald Steven Falchook, Geoffrey Shapiro, Ki Y Chung, Robert Martin Conry, David S. Hong, Judy Sing-Zan Wang, Ulrich Steidl, Loren D. Walensky, Vincent Guerlavais, Marie Payton, D. Allen Annis, Manuel Aivado, and Manish R. Patel. Phase I trial of a novel stapled peptide ALRN-6924 disrupting MDMX- and MDM2-mediated inhibition of WT p53 in patients with solid tumors and lymphomas. . *J. Clin. Oncol.*, 35(15_suppl):2505–2505, 2017.

[52] Koch R., Christie A.L., Yoshida N., Ng S.Y., Shigemori K., Morrow S.N., Van Scoyk A., Horwitz S.M., Jacobsen E., Santiago S., Ren J.-G., Guerlavais V., Annis D.A., Aivado M.A., and Weinstock D.M. In vitro and pre-clinical in vivo evidence support MDMX/MDM2 as common vulnerabilities across tp53 -wild-type t-cell lymphomas that are targetable with the α-helical p53 stapled peptide ALRN-6924. *Blood*, 130 (1), 2017.

[53] M. Payton, D. Pinchasik, A. Mehta, S. Goel, J.M. Zain, L. Sokol, E. Jacobsen, M.R. Patel, S.M. Horwitz, F. Meric-Bernstam, A. Shustov, D. Weinstock, M. Aivado, and D.A. Annis. 1039TiPPhase 2a study of a novel stapled peptide ALRN-6924 disrupting MDMX- and MDM2-mediated inhibition of wild-type TP53 in patients with peripheral t-cell lymphoma. *Ann. Oncol.*, 28(suppl_5), 2017.

[54] Aileron Therapeutics. Clinical Programs | Aileron. https://www.aileronrx.com/clinical-programs/.

[55] Fabrice Lefèvre, Marie Hélène Rémy, and Jean Michel Masson. Alanine-stretch scanning mutagenesis: A simple and efficient method to probe protein structure and function. *Nucleic Acids Res.*, 25(2):447–448, 1997.

[56] Gregory A. Weiss, Colin K. Watanabe, Alan Zhong, Audrey Goddard, and Sachdev S. Sidhu. Rapid mapping of protein functional epitopes by combinatorial alanine scanning. *Proc. Natl. Acad. Sci. U. S. A.*, 97(16):8950–8954, 2000.

[57] Elisabeth L. Humphris and Tanja Kortemme. Design of multi-specificity in protein interfaces. *PLoS Comput. Biol.*, 3(8):1591–1604, 2007.

[58] Saliha Ece Acuner Ozbabacan, Hatice Billur Engin, Attila Gursoy, and Ozlem Ke-

skin. Transient protein-protein interactions. *Protein Eng. Des. Sel.*, 24(9):635–648, 2011.

[59] George P. Smith. Filamentous fusion phage: Novel expression vectors that display cloned antigens on the virion surface. *Science (80-.).*, 228(4705):1315–1317, 1985.

[60] Greg Winter, Andrew D. Griffiths, Robert E. Hawkins, and Hennie R. Hoogenboom. Making Antibodies by Phage Display Technology. *Annu. Rev. Immunol.*, 12(1):433–455, 1994.

[61] Burckhard Seelig. MRNA display for the selection and evolution of enzymes from in vitro-translated protein libraries. *Nat. Protoc.*, 6(4):540–552, 2011.

[62] Christian Zahnd, Patrick Amstutz, and Andreas Plückthun. Ribosome display: Selecting and evolving proteins in vitro that specifically bind to a target. *Nat. Methods*, 4(3):269–279, 2007.

[63] Christopher A. Lipinski, Franco Lombardo, Beryl W. Dominy, and Paul J. Feeney. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.*, 23(1-3):3–25, 1997.

[64] R. Scott Lokey, Jennifer Liras, Spiros Liras, and Matthew P. Jacobson. Designing Molecules to Cross Biological Membranes. *Curr. Top. Med. Chem.*, 13(7):775–775, 2013.

[65] Akihiro Furukawa, Chad E. Townsend, Joshua Schwochert, Cameron R. Pye, Maria A. Bednarek, and R. Scott Lokey. Passive Membrane Permeability in Cyclic Peptomer Scaffolds Is Robust to Extensive Variation in Side Chain Functionality and Backbone Geometry. *J. Med. Chem.*, 59(20):9503–9512, 2016.

[66] Joshua Schwochert, Rushia Turner, Melissa Thang, Ray F. Berkeley, Alexandra R. Ponkey, Kelsie M. Rodriguez, Siegfried S.F. Leung, Bhagyashree Khunte, Gilles Goetz, Chris Limberakis, Amit S. Kalgutkar, Heather Eng, Michael J. Shapiro, Alan M. Mathiowetz, David A. Price, Spiros Liras, Matthew P. Jacobson, and R. Scott Lokey. Peptide to Peptoid Substitutions Increase Cell Permeability in Cyclic Hexapeptides. *Org. Lett.*, 17(12):2928–2931, 2015.

[67] William M. Hewitt, Siegfried S.F. Leung, Cameron R. Pye, Alexandra R. Ponkey, Maria Bednarek, Matthew P. Jacobson, and R. Scott Lokey. Cell-permeable cyclic peptides from synthetic libraries inspired by natural products. *J. Am. Chem. Soc.*, 137(2):715–721, 2015.

[68] Arup K. Ghose, Vellarkad N. Viswanadhan, and John J. Wendoloski. Prediction of hydrophobic (lipophilic) properties of small organic molecules using fragmental methods: An analysis of ALOGP and CLOGP methods. *J. Phys. Chem. A*, 102 (21):3762–3772, 1998.

[69] R. Scott Lokey. The Permeability Landscape around Lariat Cyclic Peptides. In *Permeability Landsc. around Lariat Cycl. Pept.*, San Diego, CA, USA, 2019. Drug Discovery Chemistry.

[70] Lisa Sercombe, Tejaswi Veerati, Fatemeh Moheimani, Sherry Y. Wu, Anil K. Sood, and Susan Hua. Advances and challenges of liposome assisted drug delivery. *Front. Pharmacol.*, 6:286, 2015.

[71] Aaron C. Anselmo and Samir Mitragotri. Nanoparticles in the clinic. *Bioeng. Transl. Med.*, 4(3):e10143, 2016.

[72] Shu jun Cao, Shuo Xu, Hui ming Wang, Yong Ling, Jiahua Dong, Rui dong Xia, and Xiang hong Sun. Nanoparticles: Oral Delivery for Protein and Peptide Drugs. *AAPS PharmSciTech*, 20(190), 2019.

[73] Sonali B. Fonseca, Mark P. Pereira, and Shana O. Kelley. Recent advances in the use of cell-penetrating peptides for medical and biological applications. *Adv. Drug Deliv. Rev.*, 61(11):953–964, 2009.

[74] Hossein Derakhshankhah and Samira Jafari. Cell penetrating peptides: A concise review with emphasis on biomedical applications. *Biomed. Pharmacother.*, 108: 1090–1096, 2018.

[75] Alan D. Frankel and Carl O. Pabo. Cellular uptake of the tat protein from human immunodeficiency virus. *Cell*, 55(6):1189–1193, 1988.

[76] André Ziegler, Pierluigi Nervi, Markus Dürrenberger, and Joachim Seelig. The cationic cell-penetrating peptide CPPTAT derived from the HIV-1 protein TAT is rapidly transported into living fibroblasts: Optical, biophysical, and metabolic evidence. *Biochemistry*, 44(1):138–148, 2005.

[77] F. Perez, A. Joliot, E. Bloch-Gallego, A. Zahraoui, A. Triller, and A. Prochiantz. Antennapedia homeobox as a signal for the cellular internalization and nuclear addressing of a small exogenous peptide. *J. Cell Sci.*, 31(2):717–722, 1992.

[78] Daniele Derossi, Alain H. Joliot, Gérard Chassaing, and Alain Prochiantz. The third helix of the Antennapedia homeodomain translocates through biological membranes. *J. Biol. Chem.*, 269(14):10444–10450, 1994.

[79] Simon W. Jones, Richard Christison, Ken Bundell, Catherine J. Voyce, Sarah M.V. Brockbank, Peter Newham, and Mark A. Lindsay. Characterisation of cell-penetrating peptide-mediated peptide delivery. *Br. J. Pharmacol.*, 145(8):1093–1102, 2005.

[80] Haoyu Tang, Lichen Yin, Kyung Hoon Kim, and Jianjun Cheng. Helical poly(arginine) mimics with superior cell-penetrating and molecular transporting properties. *Chem. Sci.*, 4(10):3839–3844, 2013.

[81] Kim Quach, Jonathan LaRochelle, Xiao Han Li, Elizabeth Rhoades, and Alanna

Schepartz. Unique arginine array improves cytosolic localization of hydrocarbon-stapled peptides. *Bioorganic Med. Chem.*, 26(6):1197–1202, 2018.

[82] Josias H. Hamman, Gill M. Enslin, and Awie F. Kotzé. Oral delivery of peptide drugs: Barriers and developments. *BioDrugs*, 19(3):165–177, 2005.

[83] Michael F. Powell, Tracy Stewart, Laszlo Otvos, Laszlo Urge, Federico C.A. Gaeta, Alessandro Sette, Thomas Arrhenius, David Thoon, Ken Soda, and Sonia M. Colon. Peptide Stability in Drug Development. II. Effect of Single Amino Acid Substitution and Glycosylation on Peptide Reactivity in Human Serum. *Pharm. Res. An Off. J. Am. Assoc. Pharm. Sci.*, 10(9):1268–1273, 1993.

[84] Regina Tugyi, Katalin Uray, Dóra Iván, Erzsébet Fellinger, Alan Perkins, and Ferenc Hudecz. Partial D-amino acid substitution: Improved enzymatic stability and preserved Ab recognition of a MUC2 epitope peptide. *Proc. Natl. Acad. Sci. U. S. A.*, 102(2):413–418, 2005.

[85] Michael Chorev and Murray Goodman. A Dozen Years of Retro-Inverso Peptidomimeticst. *Acc. Chem. Res.*, 26(5):266–273, 1993.

[86] Zhaoqianqi Feng and Bing Xu. Inspiration from the mirror: D-amino acid containing peptides in biomedical approaches. *Biomol. Concepts*, 7(3):179–187, 2016.

[87] Chong Li, Marzena Pazgier, Jing Li, Changqing Li, Min Liu, Guozhang Zou, Zhenyu Li, Jiandong Chen, Sergey G. Tarasov, Wei Yue Lu, and Wuyuan Lu. Limitations of peptide retro-inverso isomerization in molecular mimicry. *J. Biol. Chem.*, 285 (25):19572–19581, 2010.

[88] Alanod D. AlQahtani, David O'Connor, Alexander Domling, and Sayed K. Goda. Strategies for the production of long-acting therapeutics and efficient drug delivery for cancer treatment. *Biomed. Pharmacother.*, 113:108750, 2019.

[89] Daniel Steiner, Frieder W. Merz, Ivo Sonderegger, Maya Gulotti-Georgieva, Denis Villemagne, Douglas J. Phillips, Patrik Forrer, Michael T. Stumpp, Christof Zitt, and H. Kaspar Binz. Half-life extension using serum albumin-binding DARPin® domains. *Protein Eng. Des. Sel.*, 30(9):583–591, 2017.

[90] Lisa Pollaro, Sandeep Raghunathan, Julia Morales-Sanfrutos, Alessandro Angelini, Stephan Kontos, and Christian Heinis. Bicyclic peptides conjugated to an albumin-binding tag diffuse efficiently into solid tumors. *Mol. Cancer Ther.*, 14(1):151–161, 2015.

[91] Lotte Bjerre Knudsen and Jesper Lau. The discovery and development of liraglutide and semaglutide. *Front. Endocrinol. (Lausanne).*, 10(155):1–32, 2019.

[92] Ananyo A. Bhattacharya, Tim Grüne, and Stephen Curry. Crystallographic analysis reveals common modes of binding of medium and long-chain fatty acids to human serum albumin. *J. Mol. Biol.*, 303(5):721–732, 2000.

[93] Yu Heng Lau, Peterson de Andrade, Yuteng Wu, and David R Spring. Peptide stapling techniques based on different macrocyclisation chemistries. *Chem. Soc. Rev.*, 44(1):91–102, 2015.

[94] John W. Taylor. The synthesis and study of side-chain lactam-bridged peptides. *Biopolym. - Pept. Sci. Sect.*, 66(1):49–75, 2002.

[95] Mario Scrima, Alexandra Le Chevalier-Isaad, Paolo Rovero, Anna Maria Papini, Michael Chorev, and Anna Maria D'Ursi. Cu(I) catalyzed azide-alkyne intramolecular i-to-(i+4) side-chain-to-side-Chain cyclization promotes the formation of Helix-Like secondary structures. *European J. Org. Chem.*, 2010(3):446–457, 2010.

[96] Yu Heng Lau, Peterson De Andrade, Soo Tng Quah, Maxim Rossmann, Luca Laraia, Niklas Sköld, Tze Jing Sum, Pamela J.E. Rowling, Thomas L Joseph, Chandra Verma, Marko Hyvönen, Laura S Itzhaki, Ashok R Venkitaraman, Christopher J Brown, David P Lane, and David R Spring. Functionalised staple linkages for modulating the cellular activity of stapled peptides. *Chem. Sci.*, 5(5):1804–1809, 2014.

[97] Christian E Schafmeister, Julia Po, and Gregory L Verdine. An All-Hydrocarbon Cross-Linking System for Enhancing the Helicity and Metabolic Stability of Peptides. *J. Am. Chem. Soc.*, 122(24):5891–5892, 2000.

[98] Gerard J. Hilinski, Young Woo Kim, Jooyeon Hong, Peter S. Kutchukian, Charisse M. Crenshaw, Shaunna S. Berkovitch, Andrew Chang, Sihyun Ham, and Gregory L. Verdine. Stitched α-helical peptides via bis ring-closing metathesis. *J. Am. Chem. Soc.*, 136(35):12314–12322, 2014.

[99] Yu Heng Lau, Yuteng Wu, Peterson de Andrade, Warren R J D Galloway, and David R Spring. A two-component 'double-click' approach to peptide stapling. *Nat. Protoc.*, 10(4):585–94, 2015.

[100] Alexander M. Spokoyny, Yekui Zou, Jingjing J. Ling, Hongtao Yu, Yu Shan Lin, and Bradley L. Pentelute. A perfluoroaryl-cysteine S(N)Ar chemistry approach to unprotected peptide stapling. *J. Am. Chem. Soc.*, 135(16):5946–5949, 2013.

[101] J. Christopher Phelan, Nicholas J. Skelton, Andrew C. Braisted, and Robert S. McDowell. A general method for constraining short peptides to an α-helical conformation. *J. Am. Chem. Soc.*, 119(3):455–460, 1997.

[102] Wenshu Xu, Yu Heng Lau, Gerhard Fischer, Yaw Sing Tan, Anasuya Chattopadhyay, Marc De La Roche, Marko Hyvönen, Chandra Verma, David R. Spring, and Laura S. Itzhaki. Macrocyclized Extended Peptides: Inhibiting the Substrate-Recognition Domain of Tankyrase. *J. Am. Chem. Soc.*, 139(6):2245–2256, 2017.

[103] Gregory H Bird, Navid Madani, Alisa F Perry, Amy M Princiotto, Jeffrey G Supko, Xiaoying He, Evripidis Gavathiotis, Joseph G Sodroski, and Loren D Walensky. Hy-

drocarbon double-stapling remedies the proteolytic instability of a lengthy peptide therapeutic. *Proc. Natl. Acad. Sci. U. S. A.*, 107(32):14093–14098, 2010.

[104] Christopher J. Brown, Soo T. Quah, Janice Jong, Amanda M. Goh, Poh C. Chiam, Kian H. Khoo, Meng L. Choong, May A. Lee, Larisa Yurlova, Kourosh Zolghadr, Thomas L. Joseph, Chandra S. Verma, and David P. Lane. Stapled peptides with improved potency and specificity that activate p53. *ACS Chem. Biol.*, 8(3):506–512, 2013.

[105] Yong S Chang, Bradford Graves, Vincent Guerlavais, Christian Tovar, Kathryn Packman, Kwong-Him To, Karen A Olson, Kamala Kesavan, Pranoti Gangurde, Aditi Mukherjee, Theresa Baker, Krzysztof Darlak, Carl Elkin, Zoran Filipovic, Farooq Z Qureshi, Hongliang Cai, Pamela Berry, Eric Feyfant, Xiangguo E Shi, James Horstick, D Allen Annis, Anthony M Manning, Nader Fotouhi, Huw Nash, Lyubomir T Vassilev, and Tomi K Sawyer. Stapled α-helical peptide drug development: a potent dual inhibitor of MDM2 and MDMX for p53-dependent cancer therapy. *Proc. Natl. Acad. Sci. U. S. A.*, 110(36):E3445–54, 2013.

[106] Federico Bernal, Mark Wade, Marina Godes, Tina N. Davis, David G. Whitehead, Andrew L. Kung, Geoffrey M. Wahl, and Loren D. Walensky. A Stapled p53 Helix Overcomes HDMX-Mediated Suppression of p53. *Cancer Cell*, 18(5):411–422, 2010.

[107] Matteo Pellegrini, Edward M. Marcotte, and Todd O. Yeates. A fast algorithm for genome-wide analysis of proteins with repeated sequences. *Proteins Struct. Funct. Genet.*, 35(4):440–446, 1999.

[108] Andrey V. Kajava. Tandem repeats in proteins: From sequence to structure. *J. Struct. Biol.*, 179(3):279–288, 2012.

[109] Andrey V. Kajava. What curves α-solenoids? Evidence for an α-helical toroid structure of Rpn1 and Rpn2 proteins of the 26 S proteasome. *J. Biol. Chem.*, 277 (51):49791–49798, 2002.

[110] Andrey V Kajava and Alasdair C Steven. Beta-rolls, beta-helices, and other beta-solenoid proteins. *Adv. Protein Chem.*, 73:55–96, 2006.

[111] Alison Grinthal, Ivana Adamovic, Beth Weiner, Martin Karplus, and Nancy Kleckner. PR65, the HEAT-repeat scaffold of phosphatase PP2A, is an elastic connector that links force and catalysis. *Proc. Natl. Acad. Sci. U. S. A.*, 2010.

[112] J L Riffell, C J Lord, and a Ashworth. Tankyrase-targeted therapeutics: expanding opportunities in the PARP family. *Nat Rev Drug Discov*, 11(December):923–936, 2012.

[113] Patrice J. Morin. B-Catenin Signaling and Cancer. *BioEssays*, 21(12):1021–1030, 1999.

[114] Jeffrey R. Miller, Anne M. Hocking, Jeffrey D. Brown, and Randall T. Moon. Mech-

anism and function of signal transduction by the Wnt/B-catenin and Wnt/Ca2+ pathways. *Oncogene*, 18(55):7860–7872, 1999.

[115] Hiroaki Higashitsuji, Katsuhiko Itoh, Toshikazu Nagao, Simon Dawson, Kohsuke Nonoguchi, Tsuneo Kido, R. John Mayer, Shigeki Arii, and Jun Fujita. Reduced stability of retinoblastoma protein by gankyrin, an oncogenic ankyrin-repeat protein overexpressed in hepatomas. *Nat. Med.*, 6(1):96–99, 2000.

[116] Hiroaki Higashitsuji, Hisako Higashitsuji, Katsuhiko Itoh, Toshiharu Sakurai, Toshikazu Nagao, Haruhiko Sumitomo, Tomoko Masuda, Simon Dawson, Yutaka Shimada, R. John Mayer, and Jun Fujita. The oncoprotein gankyrin binds to MDM2/HDM2, enhancing ubiquitylation and degradation of p53. *Cancer Cell*, 8 (1):75–87, 2005.

[117] Hiroaki Higashitsuji, Hisako Higashitsuji, Yu Liu, Tomoko Masuda, Takanori Fujita, H. Ismail Abdel-Aziz, Supranee Kongkham, Simon Dawson, R. John Mayer, Yoshito Itoh, Toshiharu Sakurai, Katsuhiko Itoh, and Jun Fujita. The oncoprotein gankyrin interacts with RelA and suppresses NF-κB activity. *Biochem. Biophys. Res. Commun.*, 363(3):879–884, 2007.

[118] Seamus Morrone, Zhihong Cheng, Randall T Moon, Feng Cong, and Wenqing Xu. Crystal structure of a Tankyrase-Axin complex and its implications for Axin turnover and Tankyrase substrate recruitment. *Proc. Natl. Acad. Sci. U. S. A.*, 109 (5):1500–1505, 2012.

[119] Marinella G. Callow, Hoanh Tran, Lilian Phu, Ted Lau, James Lee, Wendy N. Sandoval, Peter S. Liu, Sheila Bheddah, Janet Tao, Jennie R. Lill, Jo Anne Hongo, David Davis, Donald S. Kirkpatrick, Paul Polakis, and Mike Costa. Ubiquitin ligase RNF146 regulates tankyrase and Axin to promote Wnt signaling. *PLoS One*, 6(7): e22595, 2011.

[120] Shih-Min a Huang, Yuji M Mishina, Shanming Liu, Atwood Cheung, Frank Stegmeier, Gregory a Michaud, Olga Charlat, Elizabeth Wiellette, Yue Zhang, Stephanie Wiessner, Marc Hild, Xiaoying Shi, Christopher J Wilson, Craig Mickanin, Vic Myer, Aleem Fazal, Ronald Tomlinson, Fabrizio Serluca, Wenlin Shao, Hong Cheng, Michael Shultz, Christina Rau, Markus Schirle, Judith Schlegl, Sonja Ghidelli, Stephen Fawell, Chris Lu, Daniel Curtis, Marc W Kirschner, Christoph Lengauer, Peter M Finan, John a Tallarico, Tewis Bouwmeester, Jeffery a Porter, Andreas Bauer, and Feng Cong. Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling. *Nature*, 461(7264):614–620, 2009.

[121] H L Guo, C Zhang, Q Liu, Q Li, G Lian, D Wu, X Li, W Zhang, Y Shen, Z Ye, S Y Lin, and S C Lin. The Axin/TNKS complex interacts with KIF3A and is required for insulin-stimulated GLUT4 translocation. *Cell Res*, 22(8):1246–1257, 2012.

[122] Tsung-Yin J Yeh, Juan I Sbodio, Zhi-Yang Tsun, Biao Luo, and Nai-Wen Chi. Insulin-stimulated exocytosis of GLUT4 is enhanced by IRAP and its partner tankyrase. *Biochem. J.*, 402(2):279–290, 2007.

[123] Agnes D. Berendsen and Bjorn R. Olsen. Tankyrase loses its grip on SH3BP2 in cherubism. *Cell*, 147(6):1222–1223, 2011.

[124] Noam Levaot, Oleksandr Voytyuk, Ioannis Dimitriou, Fabrice Sircoulomb, Arun Chandrakumar, Marcel Deckert, Paul M. Krzyzanowski, Andrew Scotter, Shengqing Gu, Salima Janmohamed, Feng Cong, Paul D. Simoncic, Yasuyoshi Ueki, Jose La Rose, and Robert Rottapel. Loss of Tankyrase-mediated destruction of 3BP2 is the underlying pathogenic mechanism of cherubism. *Cell*, 147(6):1324–1339, 2011.

[125] H. Seimiya. The telomeric PARP, tankyrases, as targets for cancer therapy. *Br. J. Cancer*, 94(3):341–345, 2006.

[126] Travis Eisemann, Michael McCauley, Marie-France Langelier, Kushol Gupta, Swati Roy, Gregory D. Van Duyne, and John M. Pascal. Tankyrase-1 Ankyrin Repeats Form an Adaptable Binding Platform for Targets of ADP-Ribose Modification. *Structure*, 24(10):1679–1692, 2016.

[127] Manu De Rycker, Ranga N. Venkatesan, Chao Wei, and Carolyn M. Price. Vertebrate tankyrase domain structure and sterile α motif (SAM)-mediated multimerization. *Biochem. J.*, 372(1):87–96, 2003.

[128] Manu De Rycker and Carolyn M Price. Tankyrase Polymerization Is Controlled by Its Sterile Alpha Motif and Poly(ADP-Ribose) Polymerase Domains. *Mol. Cell. Biol.*, 24(22):9802–9812, 2004.

[129] Laura Mariotti, Catherine M. Templeton, Michael Ranes, Patricia Paracuellos, Nora Cronin, Fabienne Beuron, Edward Morris, and Sebastian Guettler. Tankyrase Requires SAM Domain-Dependent Polymerization to Support Wnt-β-Catenin Signaling. *Mol. Cell*, 63(3):498–513, 2016.

[130] Juan I. Sbodio and Nai Wen Chi. Identification of a tankyrase-binding motif shared by IRAP, TAB182, and human TRF1 but not mouse TRF1: NuMA contains this RXXPDG motif and is a novel tankyrase partner. *J. Biol. Chem.*, 277(35):31887–31892, 2002.

[131] Sebastian Guettler, Jose Larose, Evangelia Petsalaki, Gerald Gish, Andy Scotter, Tony Pawson, Robert Rottapel, and Frank Sicheri. Structural basis and sequence rules for substrate recognition by tankyrase explain the basis for cherubism disease. *Cell*, 147(6):1340–1354, 2011.

[132] Zhizhi Wang, Gregory A Michaud, Zhihong Cheng, Yue Zhang, Thomas R Hinds, Erkang Fan, Feng Cong, and Wenqing Xu. Recognition of the iso-ADP-ribose moiety in poly(ADP-ribose) by WWE domains suggests a general mechanism for poly

(ADP-ribosyl)ation-dependent ubiquitination. *Genes Dev.*, 26(3):235–240, 2012.

[133] Ryan C. Dregalla, Junqing Zhou, Rupa R. Idate, Christine L R Battaglia, Howard L. Liber, and Susan M. Bailey. Regulatory roles of tankyrase 1 at telomeres and in DNA repair: Suppression of T-SCE and stabilization of DNA-pkcs. *Aging (Albany. NY).*, 2(10):691–708, 2010.

[134] Paul A. DaRosa, Rachel E Klevit, and Wenqing Xu. Structural basis for tankyrase-RNF146 interaction reveals noncanonical tankyrase-binding motifs. *Protein Sci.*, 27 (6):1057–1067, 2018.

[135] Zita Nagy, Alkmini Kalousi, Audrey Furst, Marc Koch, Benoit Fischer, and Evi Soutoglou. Tankyrases Promote Homologous Recombination and Check Point Activation in Response to DSBs. *PLoS Genet.*, 12(2):e1005791, 2016.

[136] Mi Kyung Kim. Novel insight into the function of Tankyrase. *Oncol. Lett.*, 16(6): 6895–6902, 2018.

[137] Sebastian Guettler. AXIN Shapes Tankyrase ARChitecture. *Structure*, 24(10): 1625–1627, 2016.

[138] J. M. Lee, J. A. Ledermann, and E. C. Kohn. PARP inhibitors for BRCA1/2 mutation-associated and BRCA-like malignancies. *Ann. Oncol.*, 25(1):32–40, 2014.

[139] N. V. Malyuchenko, E. Yu Kotova, O. I. Kulaeva, M. P. Kirpichnikov, and V. M. Studitskiy. PARP1 Inhibitors: Antitumor drug design. *Acta Naturae*, 7(3):27–37, 2015.

[140] Damien D'Amours, Serge Desnoyers, Icy D'Silva, and Guy G. Poirier. Poly(ADP-ribosyl)ation reactions in the regulation of nuclear functions. *Biochem. J.*, 342(2): 249–268, 1999.

[141] Elisabet Wahlberg, Tobias Karlberg, Ekaterina Kouznetsova, Natalia Markova, Antonio Macchiarulo, Ann Gerd Thorsell, Ewa Pol, Åsa Frostell, Torun Ekblad, Delal Öncü, Björn Kull, Graeme Michael Robertson, Roberto Pellicciari, Herwig Schüler, and Johan Weigelt. Family-wide chemical profiling and structural analysis of PARP and tankyrase inhibitors. *Nat. Biotechnol.*, 30(3):283–288, 2012.

[142] Christina A. Kirby, Atwood Cheung, Aleem Fazal, Michael D. Shultz, and Travis Stams. Structure of human tankyrase 1 in complex with small-molecule inhibitors PJ34 and XAV939. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.*, 68(2): 115–118, 2012.

[143] Mohit Narwal, Harikanth Venkannagari, and Lari Lehtiö. Structural basis of selective inhibition of human tankyrases. *J. Med. Chem.*, 55(3):1360–1367, 2012.

[144] Ted Lau, Emily Chan, Marinella Callow, Jo Waaler, Jason Boggs, Robert A. Blake, Steven Magnuson, Amy Sambrone, Melissa Schutten, Ron Firestein, Ondrej Machon, Vladimir Korinek, Edna Choo, Dolores Diaz, Mark Merchant, Paul Polakis,

Daniel D. Holsworth, Stefan Krauss, and Mike Costa. A novel tankyrase small-molecule inhibitor suppresses APC mutation-driven colorectal tumor growth. *Cancer Res.*, 73(10):3132–3144, 2013.

[145] Richard G. James, Kathryn C. Davidson, Katherine A. Bosch, Travis L. Biechele, Nicholas C. Robin, Russell J. Taylor, Michael B. Major, Nathan D. Camp, Kerry Fowler, Timothy J. Martins, and Randall T. Moon. WIKI4, a Novel Inhibitor of Tankyrase and Wnt/ß-Catenin Signaling. *PLoS One*, 7(12), 2012.

[146] Howard Bregman, Hakan Gunaydin, Yan Gu, Steve Schneider, Cindy Wilson, Erin F. Dimauro, and Xin Huang. Discovery of a class of novel tankyrase inhibitors that bind to both the nicotinamide pocket and the induced pocket. *J. Med. Chem.*, 56(3):1341–1345, 2013.

[147] Katie Pollock, Manjuan Liu, Mariola Zaleska, Mirco Meniconi, Mark Pfuhl, Ian Collins, and Sebastian Guettler. Fragment-based screening identifies molecules targeting the substrate-binding ankyrin repeat domains of tankyrase. *bioRxiv*, 9(1): 567446, 2019.

[148] V. Sudakin, D. Ganoth, A. Dahan, H. Heller, J. Hershko, F. C. Luca, J. V. Ruderman, and A. Hershko. The cyclosome, a large complex containing cyclin-selective ubiquitin ligase activity, targets cyclins for destruction at the end of mitosis. *Mol. Biol. Cell*, 6(2):185–198, 1995.

[149] Randall W. King, Jan Michael Peters, Stuart Tugendreich, Mark Rolfe, Philip Hieter, and Marc W. Kirschner. A 20s complex containing CDC27 and CDC16 catalyzes the mitosis-specific conjugation of ubiquitin to cyclin B. *Cell*, 81(2):279–288, 1995.

[150] Leifu Chang, Ziguo Zhang, Jing Yang, Stephen H. McLaughlin, and David Barford. Atomic structure of the APC/C and its mechanism of protein ubiquitination. *Nature*, 522(7557):450–454, 2015.

[151] Claudio Alfieri, Leifu Chang, Ziguo Zhang, Jing Yang, Sarah Maslen, Mark Skehel, and David Barford. Molecular basis of APC/C regulation by the spindle assembly checkpoint. *Nature*, 536(7617):431–436, 2016.

[152] Masaya Yamaguchi, Ryan VanderLinden, Florian Weissmann, Renping Qiao, Prakash Dube, Nicholas G. Brown, David Haselbach, Wei Zhang, Sachdev S. Sidhu, Jan Michael Peters, Holger Stark, and Brenda A. Schulman. Cryo-EM of Mitotic Checkpoint Complex-Bound APC/C Reveals Reciprocal and Conformational Regulation of Ubiquitin Ligation. *Mol. Cell*, 63(4):593–607, 2016.

[153] Nicholas G. Brown, Ryan VanderLinden, Edmond R. Watson, Florian Weissmann, Alban Ordureau, Kuen Phon Wu, Wei Zhang, Shanshan Yu, Peter Y. Mercredi, Joseph S. Harrison, Iain F. Davidson, Renping Qiao, Ying Lu, Prakash Dube,

Michael R. Brunner, Christy R R Grace, Darcie J. Miller, David Haselbach, Marc A. Jarvis, Masaya Yamaguchi, David Yanishevski, Georg Petzold, Sachdev S. Sidhu, Brian Kuhlman, Marc W. Kirschner, J. Wade Harper, Jan Michael Peters, Holger Stark, and Brenda A. Schulman. Dual RING E3 architectures regulate multiubiquitination and ubiquitin chain elongation by APC/C. *Cell*, 165(6):1440–1453, 2016.

[154] Suyang Zhang, Thomas Tischer, and David Barford. Cyclin A2 degradation during the spindle assembly checkpoint requires multiple binding modes to the APC/C. *Nat. Commun.*, 10(1), 2019.

[155] Leifu Chang and David Barford. Insights into the anaphase-promoting complex: A molecular machine that regulates mitosis. *Curr. Opin. Struct. Biol.*, 29(1):1–9, 2014.

[156] Monica C. Rodrigo-Brenni and David O. Morgan. Sequential E2s Drive Polyubiquitin Chain Assembly on APC Targets. *Cell*, 130(1):127–139, 2007.

[157] Tao Wu, Yifat Merbl, Ying Huo, Jennifer L. Gallop, Amit Tzur, and Marc W. Kirschner. UBE2S drives elongation of K11-linked ubiquitin chains by the anaphase-promoting complex. *Proc. Natl. Acad. Sci. U. S. A.*, 107(4):1355–1360, 2010.

[158] Leifu Chang, Ziguo Zhang, Jing Yang, Stephen H. McLaughlin, and David Barford. Molecular architecture and mechanism of the anaphase-promoting complex. *Nature*, 513(7518):388–393, 2014.

[159] Y. Zhang and E. Lees. Identification of an Overlapping Binding Domain on Cdc20 for Mad2 and Anaphase-Promoting Complex: Model for Spindle Checkpoint Regulation. *Mol. Cell. Biol.*, 21(15):5190–5199, 2001.

[160] Daisuke Izawa and Jonathon Pines. Mad2 and the APC/C compete for the same site on Cdc20 to ensure proper chromosome segregation. *J. Cell Biol.*, 199(1):27–37, 2012.

[161] Hartmut C. Vodermaier, Christian Gieffers, Sebastian Maurer-Stroh, Frank Eisenhaber, and Jan Michael Peters. TPR subunits of the anaphase-promoting complex mediate binding to the activator protein CDH1. *Curr. Biol.*, 13(17):1459–1468, 2003.

[162] Claudio Alfieri, Suyang Zhang, and David Barford. Visualizing the complex functions and mechanisms of the anaphase promoting complex/cyclosome (APC/C). *Open Biol.*, 7(11):170204, 2017.

[163] Zhuan Zhou, Mingjing He, Anil A. Shah, and Yong Wan. Insights into APC/C: from cellular function to diseases and therapeutics. *Cell Div.*, 11(1):9, 2016.

[164] J. L. Burton and M. J. Solomon. D box and KEN box motifs in budding yeast Hsl1p are required for APC-mediated degradation and direct binding to Cdc20p

and Cdh1p. *Genes Dev.*, 15(18):2381–2395, 2001.

[165] Jun He, William C.H. Chao, Ziguo Zhang, Jing Yang, Nora Cronin, and David Barford. Insights into degron recognition by APC/C coactivators from the structure of an Acm1-Cdh1 complex. *Mol. Cell*, 50(5):649–660, 2013.

[166] Barbara DiFiore, Norman E Davey, Anja Hagting, Daisuke Izawa, Jörg Mansfeld, Toby J Gibson, and Jonathon Pines. The ABBA Motif binds APC/C activators and is shared by APC/C substrates and regulators. *Dev. Cell*, 32(3):358–372, 2015.

[167] Laura A. Diaz-Martinez, Wei Tian, Bing Li, Ross Warrington, Luying Jia, Chad A. Brautigam, Xuelian Luo, and Hongtao Yu. The Cdc20-binding Phe box of the spindle checkpoint protein BubR1 maintains the mitotic checkpoint complex during mitosis. *J. Biol. Chem.*, 290(4):2431–2443, 2015.

[168] Norman E. Davey and David O. Morgan. Building a Regulatory Network with Short Linear Sequence Motifs: Lessons from the Degrons of the Anaphase-Promoting Complex. *Mol. Cell*, 64(1):12–23, 2016.

[169] Norman E. Davey and David O. Morgan. APC/C degron repository. http://slim.icr.ac.uk/apc/.

[170] Edgar R Kramer, Nadja Scheuringer, Alexandre V Podtelejnikov, Matthias Mann, and J M Peters. Mitotic regulation of the APC activator proteins CDC20 and CDH1. *Mol. Biol. Cell*, 11(5):1555–69, 2000.

[171] Claudine Kraft, Franz Herzog, Christian Gieffers, Karl Mechtler, Anja Hagting, Jonathon Pines, and Jan Michael Peters. Mitotic regulation of the human anaphase-promoting complex by phosphorylation. *EMBO J.*, 22(24):6598–6609, 2003.

[172] Michiel Boekhout and Rob Wolthuis. Nek2A destruction marks APC/C activation at the prophase-to-prometaphase transition by spindle-checkpoint-restricted Cdc20. *J. Cell Sci.*, 128(8):1639–1653, 2015.

[173] Daisuke Izawa and Jonathon Pines. The mitotic checkpoint complex binds a second CDC20 to inhibit active APC/C. *Nature*, 517(7536):631–634, 2015.

[174] Jörg Mansfeld, Philippe Collin, Mark O. Collins, Jyoti S. Choudhary, and Jonathon Pines. APC15 drives the turnover of MCC-CDC20 to make the spindle assembly checkpoint responsive to kinetochore attachment. *Nat. Cell Biol.*, 13(10):1234–1243, 2011.

[175] Scott A. Foster and David O. Morgan. The APC/C Subunit Mnd2/Apc15 Promotes Cdc20 Autoubiquitination and Spindle Assembly Checkpoint Inactivation. *Mol. Cell*, 47(6):921–932, 2012.

[176] Rosella Visintin, Karen Craig, Ellen S. Hwang, Susanne Prinz, Mike Tyers, and Angelika Amon. The phosphatase Cdc14 triggers mitotic exit by reversal of Cdk-dependent phosphorylation. *Mol. Cell*, 2(6):709–718, 1998.

[177] James N. Huang, Iha Park, Eric Ellingson, Laurie E. Littlepage, and David Pellman. Activity of the APC Cdh1 form of the anaphase-promoting complex persists until S phase and prevents the premature expression of Cdc20p. *J. Cell Biol.*, 154(1): 85–94, 2001.

[178] T. Kidokoro, C. Tanikawa, Y. Furukawa, T. Katagiri, Y. Nakamura, and K. Matsuda. CDC20, a potential cancer therapeutic target, is negatively regulated by p53. *Oncogene*, 27(11):1562–1571, 2008.

[179] Jia Li, Jian Zhi Gao, Jing Li Du, Zhong Xi Huang, and Li Xin Wei. Increased CDC20 expression is associated with development and progression of hepatocellular carcinoma. *Int. J. Oncol.*, 2014.

[180] Zhiwei Wang, Lixin Wan, Jiateng Zhong, Hiroyuki Inuzuka, Pengda Liu, Fazlul H. Sarkar, and Wenyi Wei. Cdc20: A Potential Novel Therapeutic Target for Cancer Treatment. *Curr. Pharm. Des.*, 19(18):3210–3214, 2013.

[181] Wen jing Wu, Kai shun Hu, De shen Wang, Zhao lei Zeng, Dong sheng Zhang, Dong liang Chen, Long Bai, and Rui hua Xu. CDC20 overexpression predicts a poor prognosis for patients with colorectal cancer. *J. Transl. Med.*, 11(1), 2013.

[182] Shengjie Wang, Borong Chen, Zhipeng Zhu, Liang Zhang, Junjie Zeng, Guoxing Xu, Gang Liu, Disheng Xiong, Qi Luo, and Zhengjie Huang. CDC20 overexpression leads to poor prognosis in solid tumors: A system review and meta-analysis. *Med. (United States)*, 97(52), 2018.

[183] Tatsuya Kato, Yataro Daigo, Masato Aragaki, Keidai Ishikawa, Masaaki Sato, and Mitsuhito Kaji. Overexpression of CDC20 predicts poor prognosis in primary non-small cell lung cancer patients. *J. Surg. Oncol.*, 106(4):423–430, 2012.

[184] Inês M.B. Moura, Maria L. Delgado, Patrícia M.A. Silva, Carlos A. Lopes, José B. do Amaral, Luís S. Monteiro, and Hassan Bousbaa. High CDC20 expression is associated with poor prognosis in oral squamous cell carcinoma. *J. Oral Pathol. Med.*, 43(3):225–231, 2014.

[185] Lixia Wang, Jinfang Zhang, Lixin Wan, Xiuxia Zhou, Zhiwei Wang, and Wenyi Wei. Targeting Cdc20 as a novel cancer therapeutic strategy. *Pharmacol. Ther.*, 151:141–151, 2015.

[186] Lin Cheng, Yu Zhou Huang, Wei Xian Chen, Liang Shi, Zhi Li, Xu Zhang, Xin Yuan Dai, Ji Fu Wei, and Qiang Ding. Cell division cycle proteinising prognostic biomarker of breast cancer. *Biosci. Rep.*, 40(5), 2020.

[187] Zhaowei Chu, Xinyue Zhang, Qingyan Li, Guanglei Hu, Christine Guo Lian, and Songmei Geng. CDC20 contributes to the development of human cutaneous squamous cell carcinoma through the Wnt/β-catenin signaling pathway. *Int. J. Oncol.*, 54(5):1534–1544, 2019.

[188] The Cancer Genome Atlas Program - National Cancer Institute. https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga.

[189] Simon A. Forbes, David Beare, Prasad Gunasekaran, Kenric Leung, Nidhi Bindal, Harry Boutselakis, Minjie Ding, Sally Bamford, Charlotte Cole, Sari Ward, Chai Yin Kok, Mingming Jia, Tisham De, Jon W. Teague, Michael R. Stratton, Ultan McDermott, and Peter J. Campbell. COSMIC: Exploring the world's knowledge of somatic mutations in human cancer. *Nucleic Acids Res.*, 43(D1):D805–D811, 2015.

[190] Thomas J. Hudson, Warwick Anderson, Axel Aretz, Anna D. Barker, Cindy Bell, Rosa R. Bernabé, M. K. Bhan, Fabien Calvo, Iiro Eerola, Daniela S. Gerhard, Alan Guttmacher, Mark Guyer, Fiona M. Hemsley, Jennifer L. Jennings, David Kerr, Peter Klatt, Patrik Kolar, Jun Kusuda, David P. Lane, Frank Laplace, Youyong Lu, Gerd Nettekoven, Brad Ozenberger, Jane Peterson, T. S. Rao, Jacques Remacle, Alan J. Schafer, Tatsuhiro Shibata, Michael R. Stratton, Joseph G. Vockley, Koichi Watanabe, Huanming Yang, Matthew M.F. Yuen, Bartha M. Knoppers, Martin Bobrow, Anne Cambon-Thomsen, Lynn G. Dressler, Stephanie O.M. Dyke, Yann Joly, Kazuto Kato, Karen L. Kennedy, Pilar Nicolás, Michael J. Parker, Emmanuelle Rial-Sebbag, Carlos M. Romeo-Casabona, Kenna M. Shaw, Susan Wallace, Georgia L. Wiesner, Nikolajs Zeps, Peter Lichter, Andrew V. Biankin, Christian Chabannon, Lynda Chin, Bruno Clément, Enrique De Alava, Françoise Degos, Martin L. Ferguson, Peter Geary, D. Neil Hayes, Amber L. Johns, Arek Kasprzyk, Hidewaki Nakagawa, Robert Penny, Miguel A. Piris, Rajiv Sarin, Aldo Scarpa, Marc Van De Vijver, P. Andrew Futreal, Hiroyuki Aburatani, Mónica Bayés, David D.L. Bowtell, Peter J. Campbel, Xavier Estivill, Sean M. Grimmond, Ivo Gut, Martin Hirst, Carlos López-Otyn, Partha Majumder, Marco Marra, John D. McPherson, Zemin Ning, Xose S. Puente, Yijun Ruan, Hendrik G. Stunnenberg, Harold Swerdlow, Victor E. Velculescu, Richard K. Wilson, Hong H. Xue, Liu Yang, Paul T. Spellman, Gary D. Bader, Paul C. Boutros, Paul Flicek, Gad Getz, Roderic Guigó, Guangwu Guo, David Haussler, Simon Heath, Tim J. Hubbard, Tao Jiang, Steven M. Jones, Qibin Li, Nuria López-Bigas, Ruibang Luo, Lakshmi Muthuswamy, B. F.Francis Ouellette, John V. Pearson, Victor Quesada, Benjamin J. Raphael, Chris Sander, Terence P. Speed, Lincoln D. Stein, Joshua M. Stuart, Jon W. Teague, Yasushi Totoki, Tatsuhiko Tsunoda, Alfonso Valencia, David A. Wheeler, Honglong Wu, Shancen Zhao, Guangyu Zhou, Mark Lathrop, Gilles Thomas, Teruhiko Yoshida, Myles Axton, Chris Gunter, Linda J. Miller, Junjun Zhang, Syed A. Haider, Jianxin Wang, Christina K. Yung, Anthony Cross, Yong Liang, Saravanamuttu Gnaneshan, Jonathan Guberman, Jack Hsu, Don R.C. Chalmers, Karl W. Hasel, Terry S.H. Kaan, William W. Lowrance, Tohru Masui, Laura Lyman Rodriguez, Catherine

Vergely, David D.L. Bowtel, Nicole Cloonan, Anna DeFazio, James R. Eshleman, Dariush Etemadmoghadam, Brooke A. Gardiner, James G. Kench, Robert L. Sutherland, Margaret A. Tempero, Nicola J. Waddell, Peter J. Wilson, Steve Gallinger, Ming Sound Tsao, Patricia A. Shaw, Gloria M. Petersen, Debabrata Mukhopadhyay, Ronald A. DePinho, Sarah Thayer, Kamran Shazand, Timothy Beck, Michelle Sam, Lee Timms, Vanessa Ballin, Jiafu Ji, Xiuqing Zhang, Feng Chen, Xueda Hu, Qi Yang, Geng Tian, Lianhai Zhang, Xiaofang Xing, Xianghong Li, Zhenggang Zhu, Yingyan Yu, Jun Yu, Jörg Tost, Paul Brennan, Ivana Holcatova, David Zaridze, Alvis Brazma, Lars Egevad, Egor Prokhortchouk, Rosamonde Elizabeth Banks, Mathias Uhlén, Juris Viksna, Fredrik Ponten, Konstantin Skryabin, P. Andrew Futrea, Ewan Birney, Ake Borg, Anne Lise Børresen-Dale, Carlos Caldas, John A. Foekens, Sancha Martin, Jorge S. Reis-Filho, Andrea L. Richardson, Christos Sotiriou, Laura Van T. Veer, Daniel Birnbaum, Hélène Blanche, Pascal Boucher, Sandrine Boyault, Jocelyne D. Masson-Jacquemier, Iris Pauporté, Xavier Pivot, Anne Vincent-Salomon, Eric Tabone, Charles Theillet, Isabelle Treilleux, Paulette Bioulac-Sage, Thomas Decaens, Franc OiseDegos, Dominique Franco, Marta Gut, Didier Samuel, Jessica Zucman-Rossi, Roland Eils, Benedikt Brors, Jan O. Korbe, Andrey Korshunov, Pablo Landgraf, Hans Lehrach, Stefan Pfister, Bernhard Radlwimmer, Guido Reifenberger, Michael D. Taylor, Christof Von Kalle, Partha P. Majumder, T. S. Rao, Paolo Pederzoli, Rita T. Lawlor, Massimo Delledonne, Alberto Bardelli, Thomas Gress, David Klimstra, Giuseppe Zamboni, Yusuke Nakamura, Satoru Miyano, Akihiro Fujimoto, Elias Campo, Silvia De Sanjosé, Emili Montserrat, Marcos González-Dyaz, Pedro Jares, Heinz Himmelbaue, Silvia Bea, Samuel Aparicio, Douglas F. Easton, Francis S. Collins, Carolyn C. Compton, Eric S. Lander, Wylie Burke, Anthony R. Green, Stanley R. Hamilton, Olli P. Kallioniemi, Timothy J. Ley, Edison T. Liu, and Brandon J. Wainwright. International network of cancer genome projects. *Nature*, 464(7291):993–998, 2010.

[191] Hayley M. Dingerdissen, John Torcivia-Rodriguez, Yu Hu, Ting Chia Chang, Raja Mazumder, and Robel Kahsay. BioMuta and BioXpress: Mutation and expression knowledgebases for cancer biomarker discovery. *Nucleic Acids Res.*, 46(D1):D1128–D1136, 2018.

[192] Patricia A.J. Muller and Karen H. Vousden. P53 mutations in cancer. *Nat. Cell Biol.*, 15(1):2–8, 2013.

[193] Wei Tian, Bing Li, Ross Warrington, Diana R Tomchick, Hongtao Yu, and Xuelian Luo. Structural analysis of human Cdc20 supports multisite degron recognition by APC/C. *Proc. Natl. Acad. Sci.*, 109(45):18419–18424, 2012.

[194] Xing Zeng, Frederic Sigoillot, Shantanu Gaur, Sungwoon Choi, Kathleen L Pfaff,

Dong-chan Oh, Nathaniel Hathaway, Nevena Dimova, Gregory D Cuny, and Randall W King. Pharmacologic Inhibition of the Anaphase-Promoting Complex Induces A Spindle Checkpoint-Dependent Mitotic Arrest in the Absence of Spindle Damage. *Cancer Cell*, 18(4):382–395, 2010.

[195] Katharine L. Sackton, Nevena Dimova, Xing Zeng, Wei Tian, Mengmeng Zhang, Timothy B Sackton, Johnathan Meaders, Kathleen L. Pfaff, Frederic Sigoillot, Hongtao Yu, Xuelian Luo, and Randall W. King. Synergistic blockade of mitotic exit by two chemical inhibitors of the APC/C. *Nature*, 514(7524):646–649, 2014.

[196] Junlong Chi, Hongchun Li, Zhuan Zhou, Javier Izquierdo-Ferrer, Yifan Xue, Cindy M Wavelet, Gary E Schiltz, Bin Zhang, Massimo Cristofanilli, Xinghua Lu, Ivet Bahar, and Yong Wan. A novel strategy to block mitotic progression for targeted therapy. *EBioMedicine*, 49:40–54, 2019.

[197] Rati Verma, Noel R. Peters, Mariapina D'Onofrio, Gregory P. Tochtrop, Kathleen M. Sakamoto, Ranjani Varadan, Mingsheng Zhang, Philip Coffino, David Fushman, Raymond J. Deshaies, and Randall W. King. Ubistatins inhibit proteasome-dependent degradation by binding the ubiquitin chain. *Science (80-.).*, 306(5693): 117–120, 2004.

[198] Xing Zeng and Randall W King. An APC/C inhibitor stabilizes cyclin B1 by prematurely terminating ubiquitination. *Nat. Chem. Biol.*, 8(4):383–392, 2012.

[199] Katherine V. Richeson, Tatyana Bodrug, Katharine L. Sackton, Masaya Yamaguchi, Joao A. Paulo, Steven P. Gygi, Brenda A. Schulman, Nicholas G. Brown, and Randall W. King. Paradoxical mitotic exit induced by a small molecule inhibitor of APC/CCdc20. *Nat. Chem. Biol.*, 16(5):546–555, 2020.

[200] Ziguo Zhang, Jing Yang, and David Barford. Recombinant expression and reconstitution of multiprotein complexes by the USER cloning method in the insect cell-baculovirus expression system. *Methods*, 95:13–25, 2016.

[201] NEB. NEB Tm Calculator. https://tmcalculator.neb.com/#!/main.

[202] Elisabeth Gasteiger, Christine Hoogland, Alexandre Gattiker, Séverine Duvaud, Marc R Wilkins, Ron D Appel, and Amos Bairoch. Protein Identification and Analysis Tools on the ExPASy Server. In *Proteomics Protoc. Handb.*, pages 571–607. 2005.

[203] Joseph E. Tropea, Scott Cherry, and David S. Waugh. Expression and purification of soluble His6-tagged TEV protease. *Methods Mol. Biol.*, 498:297–307, 2009.

[204] LLC Schrödinger and Thomas Holder. The PyMOL Molecular Graphics System, Version 2.3, 2019.

[205] Christoph Gohlke. Python Extension Packages for Windows. https://www.lfd.uci.edu/~gohlke/pythonlibs/.

[206] Martyn D. Winn, Charles C. Ballard, Kevin D. Cowtan, Eleanor J. Dodson, Paul Emsley, Phil R. Evans, Ronan M. Keegan, Eugene B. Krissinel, Andrew G.W. Leslie, Airlie McCoy, Stuart J. McNicholas, Garib N. Murshudov, Navraj S. Pannu, Elizabeth A. Potterton, Harold R. Powell, Randy J. Read, Alexei Vagin, and Keith S. Wilson. Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 67(4):235–242, 2011.

[207] B. W. Matthews. Solvent content of protein crystals. *J. Mol. Biol.*, 33(2):491–497, 1968.

[208] Katherine A. Kantardjieff and Bernhard Rupp. Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Sci.*, 12(9):1865–1871, 2003.

[209] Airlie J. McCoy, Ralf W. Grosse-Kunstleve, Paul D. Adams, Martyn D. Winn, Laurent C. Storoni, and Randy J. Read. Phaser crystallographic software. *J. Appl. Crystallogr.*, 40(4):658–674, 2007.

[210] Garib N. Murshudov, Pavol Skubák, Andrey A. Lebedev, Navraj S. Pannu, Roberto A. Steiner, Robert A. Nicholls, Martyn D. Winn, Fei Long, and Alexei A. Vagin. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 67(4):355–367, 2011.

[211] G. Bricogne, E. Blanc, M. Brandl, C. Flensbueg, P. Keller, W. Paciorek, P. Roversi, A. Sharff, O.S. Smart, C. Vonrhein, and T.O. Womack. BUSTER version 2.10.3, 2020.

[212] Oliver S. Smart, Thomas O. Womack, Claus Flensburg, Peter Keller, Włodek Paciorek, Andrew Sharff, Clemens Vonrhein, and Gérard Bricogne. Exploiting structure similarity in refinement: Automated NCS and target-structure restraints in BUSTER. *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 68(4):368–380, 2012.

[213] P. Emsley, B. Lohkamp, W. G. Scott, and K. Cowtan. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 66(4):486–501, 2010.

[214] Bernhard Lohkamp. Coot for Windows. http://bernhardcl.github.io/coot/.

[215] Yu Heng Lau and David R. Spring. Efficient synthesis of Fmoc-protected azido amino acids. *Synlett*, 2011(13):1917–1919, 2011.

[216] Zhi-xin Wang. An exact mathematical expression for describing competitive binding of two different ligands to a protein molecule. *FEBS Lett.*, 360:3–6, 1995.

[217] Xing Du, Yi Li, Yuan Ling Xia, Shi Meng Ai, Jing Liang, Peng Sang, Xing Lai Ji, and Shu Qun Liu. Insights into protein–ligand interactions: Mechanisms, models, and methods. *Int. J. Mol. Sci.*, 17(2), 2016.

[218] Sharon Min Qi Chee, Jantana Wongsantichon, Quah Soo Tng, Robert Robinson, Thomas L. Joseph, Chandra Verma, David P. Lane, Christopher J. Brown, and

Farid J. Ghadessy. Structure of a stapled peptide antagonist bound to nutlin-resistant Mdm2. *PLoS One*, 9(8):e104914, 2014.

[219] Giuseppe Papalia and David Myszka. Exploring minimal biotinylation conditions for biosensor analysis using capture chips. *Anal. Biochem.*, 403(1-2):30–35, 2010.

[220] Thermo Fisher Scientific. EZ-Link^{TM} Sulfo-NHS-LC-LC-Biotin. https://www.thermofisher.com/order/catalog/product/21338#/21338.

[221] RCSB PDB: Homepage. http://www.rcsb.org/.

[222] Helen M Berman, John Westbrook, Zukang Feng, Gary Gilliland, T N Bhat, Helge Weissig, Ilya N Shindyalov, and Philip E Bourne. The Protein Data Bank. *Nucleic Acids Res.*, 28(1), 2000.

[223] Gregory L. Verdine and Gerard J. Hilinski. Stapled peptides for intracellular drug targets. *Methods Enzymol.*, 503:3–33, 2012.

[224] Barbara Di Fiore, Claudia Wurzenberger, Norman E Davey, and Jonathon Pines. The Mitotic Checkpoint Complex Requires an Evolutionary Conserved Cassette to Bind and Inhibit Active APC/C. *Mol. Cell*, 64(6):1144–1153, 2016.

[225] William C.H. Chao, Kiran Kulkarni, Ziguo Zhang, Eric H. Kong, and David Barford. Structure of the mitotic checkpoint complex. *Nature*, 484(7393):208–213, 2012.

[226] Jeremiah J. Frye, Nicholas G. Brown, Georg Petzold, Edmond R. Watson, Christy R.R. Grace, Amanda Nourse, Marc A. Jarvis, Richard W. Kriwacki, Jan Michael Peters, Holger Stark, and Brenda A. Schulman. Electron microscopy structure of human APC/C CDH1 -EMI1 reveals multimodal mechanism of E3 ligase shutdown. *Nat. Struct. Mol. Biol.*, 20(7):827–835, 2013.

[227] Richard J. Steel, Maria A. O'Connell, and Mark Searcey. Perfluoroarene-based peptide macrocycles that inhibit the Nrf2/Keap1 interaction. *Bioorganic Med. Chem. Lett.*, 28(16):2728–2731, 2018.

[228] Mareike M Wiedmann, Yaw Sing Tan, Yuteng Wu, Shintaro Aibara, Wenshu Xu, Hannah F Sore, Chandra S Verma, Laura Itzhaki, Murray Stewart, James D Brenton, and David R Spring. Development of Cell-Permeable, Non-Helical Constrained Peptides to Target a Key Protein–Protein Interaction in Ovarian Cancer. *Angew. Chemie - Int. Ed.*, 56(2):524–529, 2017.

[229] Young Woo Kim, Peter S. Kutchukian, and Gregory L. Verdine. Introduction of all-hydrocarbon i,i + 3 staples into α-helices via ring-closing olefin metathesis. *Org. Lett.*, 12(13):3046–3049, 2010.

[230] Janet L. Burton, Yong Xiong, and Mark J. Solomon. Mechanisms of pseudosubstrate inhibition of the anaphase promoting complex by Acm1. *EMBO J.*, 30(9):1818–1829, 2011.

[231] Dan Lu, Jennifer Y Hsiao, Norman E Davey, Vanessa A. van Voorhis, Scott A

Foster, Chao Tang, and David O Morgan. Multiple mechanisms determine the order of APC/C substrate degradation in mitosis. *J. Cell Biol.*, 207(1):23–39, 2014.

[232] M. Schwab, M. Neutzner, D. Möcker, and W. Seufert. Yeast Hct1 recognizes the mitotic cyclin Clb2 and other substrates of the ubiquitin ligase APC. *EMBO J.*, 20 (18):5165–5175, 2001.

[233] Helene Labit, Kazuyuki Fujimitsu, N. Sumru Bayin, Tohru Takaki, Julian Gannon, and Hiroyuki Yamano. Dephosphorylation of Cdc20 is required for its C-box-dependent activation of the APC/C. *EMBO J.*, 31(15):3351–3362, 2012.

[234] Renping Qiao, Florian Weissmann, Masaya Yamaguchi, Nicholas G. Brown, Ryan VanderLinden, Richard Imre, Marc A. Jarvis, Michael R. Brunner, Iain F. Davidson, Gabriele Litos, David Haselbach, Karl Mechtler, Holger Stark, Brenda A. Schulman, and Jan-Michael Peters. Mechanism of APC/C ^{CDC20} activation by mitotic phosphorylation. *Proc. Natl. Acad. Sci.*, 113(19):E2570–E2578, 2016.

[235] Caroline a Schneider, Wayne S Rasband, and Kevin W Eliceiri. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods*, 9(7):671–675, 2012.

[236] Frank H Niesen, Helena Berglund, and Masoud Vedadi. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc.*, 2(9):2212–2221, 2007.

[237] Natalia J. Martinez, Rosita R. Asawa, Matthew G. Cyr, Alexey Zakharov, Daniel J. Urban, Jacob S. Roth, Eric Wallgren, Carleen Klumpp-Thomas, Nathan P. Coussens, Ganesha Rai, Shyh Ming Yang, Matthew D. Hall, Juan J. Marugan, Anton Simeonov, and Mark J. Henderson. A widely-applicable high-throughput cellular thermal shift assay (CETSA) using split Nano Luciferase. *Sci. Rep.*, 8(1), 2018.

[238] Clemens Vonrhein, Claus Flensburg, Peter Keller, Andrew Sharff, Oliver Smart, Wlodek Paciorek, Thomas Womack, and Gérard Bricogne. Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 67(4):293–302, 2011.

[239] Wolfgang Kabsch. XDS. *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 66(2):125–132, 2010.

[240] Philip Evans. Scaling and assessment of data quality. *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 62(1):72–82, 2006.

[241] Philip R. Evans and Garib N. Murshudov. How good are my data and what is the resolution? *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 69(7):1204–1214, 2013.

[242] I.J Tickle, Claus Flensburg, Peter Keller, Wlodek Paciorek, Andrew Sharff, Clemens Vonrhein, and Gérard Bricogne. STARANISO, 2019.

[243] Global Phasing. Grade Web Server. http://grade.globalphasing.org, 2019.

[244] Jessica Perrin, Thilo Werner, Nils Kurzawa, Dorothee D Childs, Mathias Kalxdorf, Anna Rutkowska, Daniel Poeckel, Daniel C Sevin, Eugenia Stonehouse, Katrin Strohmer, Bianca Heller, Douglas Thomson, Johanna Vappiani, Jana Krause, H Christian Eberl, Christina E Rau, Holger Franken, Wolfgang Huber, Maria Faelth-Savitski, Mikhail M Savitski, Marcus Bantscheff, and Giovanna Bergamini. Proteome thermal stability reflects organ physiology and identifies drug-target engagement in vivo. *bioRxiv*, page 500306, 2018.

[245] Daumantas Matulis, James K Kranz, F Raymond Salemme, and Matthew J Todd. Thermodynamic stability of carbonic anhydrase: Measurements of binding affinity and stoichiometry using thermofluor. *Biochemistry*, 44(13):5258–5266, 2005.

[246] Nan Bai, Heinrich Roder, Alex Dickson, and John Karanicolas. Isothermal Analysis of ThermoFluor Data can readily provide Quantitative Binding Affinities. *Sci. Rep.*, 9(1), 2019.

[247] Justin Hall. A simple model for determining affinity from irreversible thermal shifts. *Protein Sci.*, 28(10):1880–1887, 2019.

[248] J. M. Treherne and J. M. Young. Temperature-dependence of the kinetics of the binding of [3H]-(+)-N-methyl-4-methyldiphenhydramine to the histamine H1-receptor: comparison with the kinetics of [3H]-mepyramine. *Br. J. Pharmacol.*, 94 (3):811–822, 1988.

[249] Matthieu Schapira, Mike Tyers, Maricel Torrent, and Cheryl H. Arrowsmith. WD40 repeat domain proteins: A novel target class? *Nat. Rev. Drug Discov.*, 16(11):773–786, 2017.

[250] Christopher W. Carroll, Maria Enquist-Newman, and David O. Morgan. The APC subunit Doc1 promotes recognition of the substrate destruction box. *Curr. Biol.*, 15(1):11–18, 2005.

[251] Bettina A. Buschhorn, Georg Petzold, Marta Galova, Prakash Dube, Claudine Kraft, Franz Herzog, Holger Stark, and Jan Michael Peters. Substrate binding on the APC/C occurs between the coactivator Cdh1 and the processivity factor Doc1. *Nat. Struct. Mol. Biol.*, 18(1):6–14, 2011.

[252] Paula C.A. Da Fonseca, Eric H. Kong, Ziguo Zhang, Anne Schreiber, Mark A. Williams, Edward P. Morris, and David Barford. Structures of APC/CCdh1 with substrates identify Cdh1 and Apc10 as the D-box co-receptor. *Nature*, 470(7333): 274–280, 2011.

[253] Xinxian Qiao, Liyong Zhang, Armin M. Gamper, Takeo Fujita, and Yong Wan. APC/C-Cdh1: From cell cycle to cellular differentiation and genomic integrity. *Cell Cycle*, 9(19):3904–3912, 2010.

[254] Mikhail M. Savitski, Friedrich B.M. Reinhard, Holger Franken, Thilo Werner,

Maria Fälth Savitski, Dirk Eberhard, Daniel Martinez Molina, Rozbeh Jafari, Rebecca Bakszt Dovega, Snsan Klaeger, Bernhard Kuster, Pär Nordlund, Marcus Bantscheff, and Gerard Drewes. Tracking cancer drugs in living cells by thermal profiling of the proteome. *Science (80-.).*, 346(6205), 2014.

[255] André Mateus, Tomi A. Määttä, and Mikhail M. Savitski. Thermal proteome profiling: Unbiased assessment of protein state through heat-induced stability changes. *Proteome Sci.*, 15(1), 2017.

[256] Tomas Friman. Mass spectrometry-based Cellular Thermal Shift Assay (CETSA®) for target deconvolution in phenotypic drug discovery. *Bioorganic Med. Chem.*, 28 (1), 2020.

[257] Xiuyun Sun, Hongying Gao, Yiqing Yang, Ming He, Yue Wu, Yugang Song, Yan Tong, and Yu Rao. Protacs: Great opportunities for academia and industry. *Signal Transduct. Target. Ther.*, 4(1):1–33, 2019.

[258] Jiraborrirak Charoenpattarapreeda, Yaw Sing Tan, Jessica Iegre, Stephen J. Walsh, Elaine Fowler, Rohan S. Eapen, Yuteng Wu, Hannah F. Sore, Chandra S. Verma, Laura Itzhaki, and David R. Spring. Targeted covalent inhibitors of MDM2 using electrophile-bearing stapled peptides. *Chem. Commun.*, 55(55):7914–7917, 2019.